

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 17:17:17 ; Search time 36 Seconds

(without alignments)
1477.796 Million cell updates/sec

Title: US-09-900-425a-2

Perfect score: 7500
Sequence: 1 MMGGNTCHRMSPHGRGCPD.....MKWERHOREPDEDEDIKK 1374

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7471	99.6	1374	1 RNC_HUMAN	09nrr4 homo sapien
2	1448	19.3	1086	1 RNC_CAEEL	001326 caenorhabd
3	315	4.2	249	1 RNC_BACSU	P51833 bacillus su
4	285.5	3.8	225	1 RNC_VIBCH	09kpb2 vibrio chol
5	282	3.8	225	1 RNC_PASMU	P57805 Pasteurella
6	282	3.8	263	1 RNC_BACMD	09k405 bacillus ha
7	271.5	3.6	620	1 EXTN_TOBAC	P13583 nicotiana t
8	270	3.6	226	1 RNC_BUCAI	P57346 buchnera ap
9	268.5	3.6	231	1 RNC_LACIA	09chd0 lactococcus
10	260	3.5	226	1 RNC_ECOLI	P05797 escherichia
11	260	3.5	1832	1 ZAP3_HUMAN	P49750 homo sapien
12	259.5	3.5	555	1 RNC_ECORI	09f9q6 chlamydomon
13	259.5	3.5	1386	1 RNC_ECORI	09f9q6 chlamydomon
14	259.5	3.5	1638	1 ZAP3_MOUSE	09f9q6 chlamydomon
15	258	3.4	226	1 RNC_SALTY	P25439 drosophila
16	258	3.4	231	1 RNC_SYNY3	056056 salmonella
17	257.5	3.4	272	1 RNC_SYNY3	P74368 synecocyst
18	256	3.4	272	1 RNC_STRCO	09zbn7 streptomyce
19	253.5	3.4	246	1 RNC_BOBBU	051448 borrelia bu
20	251.5	3.4	1048	1 RNC_HUMAN	063627 rattus norv
21	245	3.3	239	1 RNC_HELPY	P56118 hellicobacte
22	244.5	3.3	1157	1 RNC_HUMAN	095104 homo sapien
23	242.5	3.2	248	1 RNC_SPCIC	053844 spliroplasma
24	242	3.2	227	1 RNC_HAEIN	P44441 haemophilus
25	241	3.2	239	1 RNC_HELPJ	09z1b2 hellicobacte
26	234	3.1	1362	1 BRD4_HUMAN	060885 homo sapien
27	233	3.1	1183	1 DRPL_RAT	P54558 rattus norv
28	233	3.1	229	1 RNC_PSEAE	09kcx9 pseudomonas
29	232.5	3.1	232	1 RNC_CHLMU	09p977 chlamydia m
30	230.5	3.1	1647	1 SN24_HUMAN	P51532 homo sapien
31	229.5	3.1	254	1 RNC_TREPA	083787 treponema p
32	229.5	3.1	424	1 S3B4_HUMAN	015427 homo sapien
33	228.5	3.0	487	1 EBN2_EBV	P12978 Epstein-bar
			237	1 RNC_CHLUPN	09z9c7 chlamydia p

34	225.5	3.0	225	1 RNC_RICPR	09a231 rickettsia
35	225	3.0	231	1 RNC_CHLTR	084299 chlamydia t
36	223.5	3.0	238	1 RNC_MYCLE	069469 mycobacteri
37	223.5	3.0	497	1 MAS2_HUMAN	097645 homo sapien
38	223	3.0	233	1 RNC_COXBU	P51837 coxiella bu
39	222.5	3.0	963	1 YQ36_CAEEL	P09457 caenorhabd
40	222	3.0	485	1 SSGP_VOLCA	P21997 volvox cart
41	221	2.9	1509	1 GSR1_HUMAN	09nzm4 homo sapien
42	220.5	2.9	1185	1 DRPL_HUMAN	P04529 homo sapien
43	218	2.9	306	1 EXTN_DAUCA	P05599 daucus caro
44	218	2.9	503	1 WARP_HUMAN	043516 homo sapien
45	218	2.9	1157	1 Y182_HUMAN	Q14687 homo sapien

ALIGNMENTS

RESULT 1
ID RNC_HUMAN STANDARD: PRT; 1374 AA.
AC Q9NRR4; Q9NRR4; Q9Y2V9; Q9Y4Y0.
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ribonuclease III (EC 3.1.26.3) (Rnase III) (P241).
GN RNASE3L OR RN3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=20538440; PubMed=10948199;
RA Wu H., Xu H., Miraglia L.J., Crooke S.T.;
RT "Human Rnase III is a 160-kDa protein involved in preribosomal RNA
RT processing.";
RT J. Biol. Chem. 275:36957-36965(2000).
RN [2]
RP SEQUENCE OF 166-613 FROM N.A. (ISOFORM 2).
RX TISSUE=Colon;
RA MEDLINE=20431278; PubMed=10976766;
RA Gunther M., Lathier M., Brison O.;
RT "A set of proteins interacting with transcription factor Spl
RT identified in a two-hybrid screening.";
RT Mol. Cell. Biochem. 210:131-142(2000).
RN [3]
RP SEQUENCE OF 603-1374 FROM N.A.
RX TISSUE=Embryo;
RA Issigal T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Nagatsuma M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Negahari K., Masuko Y.,
RA Niomiya K., Iwayanagi T.;
RT "NEO human cDNA sequencing project";
RT Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE OF 706-1374 FROM N.A.
RX TISSUE=Aorta;
RA Wei Y.D., Ding J.F., Xiong H., Zhou Y., Liew C.C.;
RT Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
RN [5]
RP FUNCTION: Involved in pre-rRNA processing. Cleaves double-strand
RN and does not cleave single-strand RNA.
CC - CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC - SUBUNIT: Interacts with Spl.
CC - SUBCELLULAR LOCATION: Nuclear. A fraction is translocated to the
CC nucleolus during the S phase of the cell cycle.
CC - ALTERNATIVE PRODUCTS: 2 isoforms, 1 (shown here) and 2; may be
CC produced by alternative splicing.
CC - TISSUE SPECIFICITY: Ubiquitous.
CC - SIMILARITY: CONTAINS 1 DBRM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.

```
CC -1- SIMILARITY: CONTAINS 2 RNASE III DOMAINS.
CC -1- CAUTION: Ref.3 sequence differs from that shown due to a
CC frameshift in position 775.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF199011; AAF80558.1; -
DR EMBL: AJ242976; CAB45133.1; -
DR EMBL: AK001121; BAA91511.1; ALT_INIT.
DR EMBL: AF116910; AAD29637.1; ALT_FRAME.
DR InterPro: IPR001159; DS_RBD.
DR InterPro: IPR000999; RNase_3.
DR Pfam: PF00035; dsrm.1.
DR Pfam: PF00636; Ribonuclease_3; 2.
DR SMART: SM00358; DSRM.1.
DR SMART: SM00535; RIBOC.2.
DR PROSITE: PS50137; DS_RBD.1.
DR PROSITE: PS00517; RNase_3_1; 2.
DR PROSITE: PS50142; RNase_3_2; 2.
KW Ribosome biogenesis; Hydrolyase; Nuclease; Endonuclease; Repeat;
KW RNA-binding; Nuclear protein; Alternative splicing.
FT DOMAIN 1 212 PRO-RICH.
FT DOMAIN 219 316 ARG-RICH.
FT DOMAIN 876 1056 RNase III 1.
FT DOMAIN 1107 1233 RNase III 2.
FT DOMAIN 1260 1334 DRBM.
FT VARSPLIC 285 353 REPERERHRHRDNRSPLESYKKEYKRSGRSYGLSVPE
PAGCPRLGELIKNTDSMAPELVN -> S (IN
ISOGORM -> RERERTSLE (IN REF. 2).
FT CONFLICT 166 174 YQYPRGYSH -> RERERTSLE (IN REF. 2).
FT CONFLICT 612 612 L -> P (IN REF. 2).
FT CONFLICT 1020 1020 R -> P (IN REF. 1).
FT CONFLICT 1230 1230 I -> T (IN REF. 1).
SQ SEQUENCE 1374 AA; 159315 MW; ED6FDEA09F3B8092 CRC64;

Query Match 99.6%; Score 7471; DB 1; Length 1374;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1370; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 421 YSSDPMDQVGDSTVVGTSRLDLYDKFEEELGSRQEKAAKAPWEPPTKTLDEDLSS 480
DB 421 YSSDPMDQVGDSTVVGTSRLDLYDKFEEELGSRQEKAAKAPWEPPTKTLDEDLSS 480
QY 481 SESECSDESDSTSSSSDSVPVYIAEIKRKAHPRLDELWYNDPGMNGPLCKGSA 540
DB 481 SESECSDESDSTSSSSDSVPVYIAEIKRKAHPRLDELWYNDPGMNGPLCKGSA 540
QY 541 KARTGRHSIYGEBAIKPCRPMTNAGRLFEYRTVSPPTNFDLPRTVLEYDDEHY 600
DB 541 KARTGRHSIYGEBAIKPCRPMTNAGRLFEYRTVSPPTNFDLPRTVLEYDDEHY 600
QY 601 FEESMFAHAPLTNIPCKYIRNIDYTHFTEEMAPENPCVKGLELSLFLPROLLEY 660
DB 601 FEESMFAHAPLTNIPCKYIRNIDYTHFTEEMAPENPCVKGLELSLFLPROLLEY 660
QY 661 DWNLKGPLFEDSPCCRFHFMRFRVFLPDGKEVLSMHQILLYLIRCSKALVPEEETA 720
DB 661 DWNLKGPLFEDSPCCRFHFMRFRVFLPDGKEVLSMHQILLYLIRCSKALVPEEETA 720
QY 721 NMLQWEELEWQYAECKGMIVTNPGRSPSSVRIDLDREOFNDVITPTIIVHFGIRPA 780
DB 721 NMLQWEELEWQYAECKGMIVTNPGRSPSSVRIDLDREOFNDVITPTIIVHFGIRPA 780
QY 781 QLSYAGDPQYQIKLMSYIVKLRLHLANSPKYKQTDOKLAOREBALOKIKQKNTMREYTV 840
DB 781 QLSYAGDPQYQIKLMSYIVKLRLHLANSPKYKQTDOKLAOREBALOKIKQKNTMREYTV 840
QY 841 ELSQSGFWKGIQSDVOQHMMMLPYLTHNIRYHQCMLNHDKLGYPFODRCILQILMTNP 900
DB 841 ELSQSGFWKGIQSDVOQHMMMLPYLTHNIRYHQCMLNHDKLGYPFODRCILQILMTNP 900
QY 901 SHHLNFGMNDHARNSLSNCGIRQPKYGDKRKVVHMMRRKGINTLINMSRLGDDPTPS 960
DB 901 SHHLNFGMNDHARNSLSNCGIRQPKYGDKRKVVHMMRRKGINTLINMSRLGDDPTPS 960
QY 961 RINNEBLEFLGDAVVEFLSYHLTYLFPSEEGGLATYTAIVQNHMLAKKLELDR 1020
DB 961 RINNEBLEFLGDAVVEFLSYHLTYLFPSEEGGLATYTAIVQNHMLAKKLELDR 1020
QY 1021 FMLYAHGPDLCRESDDLHMANCFEALIGAVYEGSLSEAKOLFGRLLFNDPDLREVMIN 1080
DB 1021 FMLYAHGPDLCRESDDLHMANCFEALIGAVYEGSLSEAKOLFGRLLFNDPDLREVMIN 1080
QY 1081 YPLHPIQLOEPNDRQILETSPLYQKLTFFEEAIGVLETHVRLLAFTLRVGFNHLTL 1140
DB 1081 YPLHPIQLOEPNDRQILETSPLYQKLTFFEEAIGVLETHVRLLAFTLRVGFNHLTL 1140
QY 1141 GHNRBMFLGDSIMQIVATYELFIHPDNHGHITLRSLSVNNRQAKVAEELGMOEYA 1200
DB 1141 GHNRBMFLGDSIMQIVATYELFIHPDNHGHITLRSLSVNNRQAKVAEELGMOEYA 1200
QY 1201 ITNDKTRPVGLTKTLADLLESLIYALYTDKIDLEVYHTPMNCFPRLKEFTLINDWMD 1260
DB 1201 ITNDKTRPVGLTKTLADLLESLIYALYTDKIDLEVYHTPMNCFPRLKEFTLINDWMD 1260
QY 1261 PKSQLQCCILTRTEGKEPDIPLYKLTQTVGSHARTYVAAYFEKGRISCGKPSIQQA 1320
DB 1261 PKSQLQCCILTRTEGKEPDIPLYKLTQTVGSHARTYVAAYFEKGRISCGKPSIQQA 1320
QY 1321 EMGAAMALEKYNPQAHOKRFLGKTYROELKEMKREHOREDEPDEDIKK 1374
DB 1321 EMGAAMALEKYNPQAHOKRFLGKTYROELKEMKREHOREDEPDEDIKK 1374

RESULT 2
RNC_CAEEL STANDARD; PRT; 1086 AA.
ID RNC_CAEEL
AC 001326; 001327; Q90908;
DT 01-NOV-1997 (Ref. 35, Created)
DT 01-MAR-2002 (Ref. 41, Last sequence update)
DT 01-MAR-2002 (Ref. 41, Last annotation update)
DE Ribonuclease III (EC 3.1.26.3) (RNase III).
```

```

GN F26E4.10 OR F26E4.13.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_taxid=6239;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Lighting J.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP REVISIONS.
RA Lighting J.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE OF 227-1086 FROM N.A.
RX MEDLINE=20179701; Pubmed=10713462;
RA Filippov V., Solovjev V., Filippova M., Gill S.S.;
RT "A novel type of RNase III family proteins in eukaryotes.";
RL Gene 245:213-221(2000).
CC -1- FUNCTION: Involved in pre-rRNA processing. Cleaves double-strand
CC RNA and does not cleave single-strand RNA (By similarity).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphonooxester.
CC -1- SUBCELLULAR LOCATION: Nucleus (Probable).
CC -1- SIMILARITY: CONTAINS 1 DNM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 RNASE III DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; Z81070; CAB03006.3; -.
DR EMBL; AF160248; AAD45518.1; -.
DR WormPep; F26E4.10; CE09694.
DR InterPro; IPR001159; DS_RBD.
DR InterPro; IPR000999; RNase_3.
DR Pfam; PF000035; dsrm; 1.
DR Pfam; PF00636; Ribonuclease_3; 2.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00535; RIBOC; 2.
DR PROSITE; PS50137; DS_RBD; 1.
DR PROSITE; PS00517; RNASE_3_1; 2.
DR PROSITE; PS00142; RNASE_3_2; 2.
KW Ribosome biogenesis; Hydrolase; Nuclease; Endonuclease; Repeat;
KW RNA-binding; Nuclear protein.
FT DOMAIN 607 781 RNASE III 1.
FT DOMAIN 833 957 RNASE III 2.
FT DOMAIN 984 1059 DRBM.
SQ SEQUENCE 1086 AA; 125333 MW; 4AA78120F8FB8CRC64;

Query Match 19.3%, Score 1448; DB 1; Length 1086;
Best Local Similarity 31.2%; Pred. No. 5.7e-65;
Matches 364; Conservative 206; Mismatches 431; Indels 166; Gaps 31.

OY 277 RHRSTERSREERHHRDRNRSPSLRSYKKEKKRSGRSTGLSVPEPACTFELP-G 335
DB 14 KHKRRARRKKYQKEYOERKHEMMOOLGRRFONOPSTSSAPPTVEKIRLPTRESTALPFG 73
OY 336 EIKKNTDS-----WAPPLETIVNHRSPSRERKKRARWEKEDRWSDNOSGKDKNYTSIK 388
DB 74 DSPRTENDYETNYMIDP--VSTHSALIK-----SNRWYIK 110
OY 389 EKEPEETMPDXNEEEHEELK-----PWIRCTHS----- 418
DB 111 AEEAEKMYMIKAKSTTSKIDOFQTKILEYVTKRRLQADVYIIHPCHSMGKRTPKX 170
OY 419 ---ENYYSDDMDVGDSTVYGTSLRLDLYDKFEEELGSRDEKAKAARP--WEPPKTKL 473

```

Dd	171	GGDSFTASVSDSDSNS-----QDASISSEPTNQAPEADT	208
Oy	474	DEDESSESECEDESDSTCS-----SSDSSEYFDVIAIKRKKAHPDRH	519
Dd	209	GE-----VDEKOTCSNRNRNOQRKARLRANFEKEQOLITAKKIDRKKTPNGIH	258
Oy	520	DELYANDBPGMNDPRLCKSKAKARNGIRISITRPGEAIRKPCRPMTNNAGRLEFHY--RT	577
Dd	259	PDISEFNEGLOINBEPRECRPEPIKTCGLKIGUYAGDKADCK--KSNENLHYTLRYT	316
Oy	578	VSPPTNFILDRPVILEVDHEIYIEGSGMFANHP-----TNPLCKVIFNFIDYIHFIE	633
Dd	317	PLPSENOL--YRTIMALNGEPEFEFEGSLTTHAPRLDCKTRABICK--YSDYEROLYE	371
Oy	634	EMMP-EMFVCYKGLBFLSLFRLDLELYDNLKGRLPDESPCCSPRPHMPRFVFLPDG	692
Dd	372	EFMDECPEDPCDMLPEYIFHEIFEMLDRLKPHIPDSVESCPIHIMPRFYQ--TKDD	430
Oy	693	GKVELSHQOILLYL--RCSKALVPEEETIANMLOMEELEMOKVAEESKGIATVNPCTKSS	751
Dd	431	LVOLMSKTYVALATPSTSGSEISMPEDVYNLCAQIDOFTRMTSKHOSIVLNTFKPSA	490
Oy	752	VRIOOLDRECPNPVITFPIIYFHNGIRPAOLSVAGROPOKLMKSVYKRLHLANSPYK	811
Dd	491	IRAWFERDEDEKKEV-----YVHNHAIQAQYTTAISLPRALFEKTLINKKIQ--EKSSGY	545
Oy	812	QTDOKLOAREALOKIROKNTKREYTVELSSOGFWKTCGIRSVCQAHMMLPVLTNHR	871
Dd	546	NKDEEK---TKNELEHLKRENRSAARNLKLREPVAGLETGLKPDVAHVMTILCHNHR	602
Oy	872	YHOSLMLDKLIGTFPDRCLOLDMATHPSHNLNFGMNDPHARNSLSNGCIQRPYTG--D	929
Dd	603	YNFSLDVEEYIEKFNDRVIELALMHSFSKSHGYETPIDHKNMTITNGYRR--KYGAD	661
Oy	930	RKYHNHMRKKGINTLITNISRLOGDPPRSRINNHRELEFGLDAVEFLTVNHYLYEP	989
Dd	662	KR-----EKKRVAGINSLEINIKGTSGGR-----TLNHERLEYLADAVELVSHNLXMTL	714
Oy	990	SLEBGLATYRTATVONQHLMLAKKLELDRFMLYVAGPDLCSRESDLRHMANCFEALIG	1049
Dd	715	HNFEGLATYRTALYQNRNLATLAKNCRIDEMLOQSHGADLLNVAEFGKHALANFEAVMA	774
Oy	1050	AVYLEGSLSEAKOLFGLLF--NDPDLREVMALYPLHPLDLOEPNDRLOLIEFSPVLOKIT	1108
Dd	775	AIYLDGGLAPCDVIFYSAMGHOQPVLKEKMDHINEHELKREDPGGRDLSFTTPTLSFPH	834
Oy	1109	EFEELAGIYFNHVALRAFTLRTYVGNHNLITGHNRMFEGDSIMOLATEYLTLRPP	1166
Dd	835	ALBERLQIOPNNITLAKKATFRNIPNNDLTKHNOBLMIGDSYQLOLVDSOFLTRRPPY	894
Oy	1169	HNHEGLTLRSSLVNNRTOAKVAEBELGOMOFYATITNDKTRPV---GLARTLTADLLESPI	1225
Dd	895	HNHEGMSLRLRSTSLVSNQTOVAVDDDGLFTBFYV-----KAPKTPBLKAKDADLVEAFI	949
Oy	1226	AALYTKDLDEYVHTFMNVVCFPRPKLEPILNQMDNPKSOLQOCCITLRL--TEKEEPDILY	1288
Dd	950	GALYVDKRIEHCRAFIIVICCPRIKHEIESEKKNDAKSHLOOWCIAMRDPSSSEBDMPEY	1009
Oy	1285	KTLDTGVGSHARTVTVVYVFKGEIIGGKQPSILOQAEKMAADALEKYFPMQAOQKPI	1344
Dd	1010	RYLDTIEGPTNRRJFKIIVYTKGRKLASABNSHNKALRYVALATA--NLESMSFK---	1066
Oy	1345	GRKYROELKEMRERHOREPEPDETJ	1371
Dd	1065	-----MKAKNMSWQNNRRRLLEDTSD	1086
RESULT 3			
RNC_BACSU			
ID	RNC_BACSU	STANDARD:	PRT: 249 AA.
AC	PS1833: 031734:		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	30-MAY-2000 (Rel. 39, Last sequence update)		

```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonuclease III (EC 3.1.26.3) (RNase III).
GN RNC OR RNCS.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96257247; PubMed=8654983;
RA Oguro A., Kakeshita H., Takamatsu H., Nakamura K., Yamane K.;
RT "The effect of Srb, a homologue of the mammalian SRP receptor alpha-
RT subunit, on Bacillus subtilis growth and protein translocation.";
RT Gene 172:17-24(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DIGESTS DOUBLE-STRANDED RNA. INVOLVED IN THE PROCESSING
CC OF RIBOSOMAL RNA PRECURSORS AND OF SOME MRNAS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: CONTAINS 1 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 RNASE III DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D64116; BAA10976.1; -
DR EMBL: Z99112; CAB13466.1; -
DR Subtilast; BG11537; rnc.
DR InterPro: IPR001159; DS_RBD.
DR InterPro: IPR000999; RNase_3.
DR Pfam: PF00035; dsrm; 1.
DR Pfam: PF00636; Ribonuclease_3; 1.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00353; RIBOC; 1.
DR PROSITE; PS00137; DS_RBD; 1.
DR PROSITE; PS00517; RNASE_3_1; 1.
DR PROSITE; PS0142; RNASE_3_2; 1.
DR Hydroxylase; Nuclease; Endonuclease; RNA-binding; Complete proteome.
KW DOMAIN
FT DOMAIN 20 149 DRBM.
FT CONFLICT 79 79 A -> P (IN REF. 1).
FT CONFLICT 171 171 S -> P (IN REF. 1).
FT CONFLICT 243 243 Q -> E (IN REF. 1).
FT CONFLICT 249 249 Q -> Q (IN REF. 1).
SQ SEQUENCE 249 AA; 28425 MW; 304F3B0B5B7CCDC CRC64;

Query Match
Best Local Similarity 4.2%; Score 315; DB 1; Length 249;
Matches 85; Conservative 46; Mismatches 96; Indels 18; Gaps 7;

QY 1104 LQKLTFFERAGIVFTVHLLARAFRTVGFNHLTLGH--NORMEFLGDSIMQVATEY 1161
DB 17 VQGFKEFORISVHQNQNELLAQAFTHSSVYVNEHKKRYEDERERLEFGDAVLELTISRF 76
QY 1162 LFIHPDDHGHGHTLRSLSVNNRTQAKVAEELGMOEATITNDKTKRPVG-LRTFTLADL 1220
DB 77 LFAKYPAWSEGDLTKLRAIVCEPSIVSLAHLSTFGDVLTLKGHEMTGGRKRPALADV 136
QY 1221 LRSFIALATDLDLEVYHFNWVCFEPRIKEFILNODMNDPKSOLQCCITLRTGK--- 1277
DB 137 FEFATGALYLDGLFESVFLKVVFPKINDGAFSHVM--DFRSQLOE---YVORDGKSL 192

```

```

QY 1278 EPDIPYKTLQTVGSPHARTYVAVYFKGBRIGCGKSPISQOAEKAMDALEKTNFPM 1337
DB 193 E-----YKISNEKGPANHREFEIVSLKGEPLGVGNRSKKEAEQHAQOALAKL---OK 244
QY 1338 AHOKR 1342
DB 245 HHTRQ 249

RESULT 4
RNC_VIRCH STANDARD; PRT; 225 AA.
ID NCBI_VIRCH
AC Q9KPB2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonuclease III (EC 3.1.26.3) (RNase III).
GN RNC OR VC2461.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Haddadson J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uitterlind T., Fleischmann R.D., Nierman W.C., White O.,
RA Fraser C.M., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RT "RNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RT Nature 406:477-483(2000).
CC -1- FUNCTION: DIGESTS DOUBLE-STRANDED RNA. INVOLVED IN THE PROCESSING
CC OF RIBOSOMAL RNA PRECURSORS AND OF SOME MRNAS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: CONTAINS 1 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 RNASE III DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE004316; AAF95603.1; -
DR TIGR: VC2461; -
DR InterPro: IPR001159; DS_RBD.
DR InterPro: IPR000999; RNase_3.
DR Pfam: PF00035; dsrm; 1.
DR Pfam: PF00636; Ribonuclease_3; 1.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00353; RIBOC; 1.
DR PROSITE; PS00137; DS_RBD; 1.
DR PROSITE; PS00517; RNASE_3_1; 1.
DR PROSITE; PS0142; RNASE_3_2; 1.
DR Hydroxylase; Nuclease; Endonuclease; RNA-binding; Complete proteome.
KW DOMAIN
FT DOMAIN 5 127 RNASE III.
FT DOMAIN 154 224 DRBM.
SQ SEQUENCE 225 AA; 25010 MW; 692PEEF4CS80990C CRC64;

Query Match
Best Local Similarity 3.8%; Score 285.5; DB 1; Length 225;
Matches 75; Conservative 42; Mismatches 101; Indels 15; Gaps 5;

QY 1101 SPVIOKLTFFERAGIVFTVHLLARAFRTVGFNHLTLGHNRMEFLGDSIMQVATEY 1160
DB 1101 SPVIOKLTFFERAGIVFTVHLLARAFRTVGFNHLTLGHNRMEFLGDSIMQVATEY 1160

```

Qy 1165 HFDPHGHGHTLLRSSVYNNRTOQAAYAEELGNQEVAITNDKTRPVGLRTKT--LADLLES 1223
|| : || : |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Dd 58 QEPFCNNGEISNRRAATLVREPRLARKEFLAGFDYLSLGPSELSSGGFRRSIIADIQVEA 117

Qy 1224 FIALIYDKLEEVHFMMNVCFEPRRIKETPLILNQDWNDRKSOLQCCUCLTLREGEKPDIPL 1283
| :: |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Dd 118 IIGALISDSLDLATTTIKIVQHMYOAKLOQRPDDNKDXPKTKILO- ----YLGGRRLPPIPT 172
| :: |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Qy 1284 YKTIQTVPSPHARTYVVAVFYFG-EIRICCGCKPSIQDAEMGCAMDAALE 1330
| :: |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:
Dd 173 VYNVEIGEHAHCOTFEVCYEVRKNIDRTFMSGASRRKADEGAAREKILO 220

RESULT 6
RNC_BACHD STANDARD; PRT; 263 AA.
ID RNC_BACHD
AC G9KA05;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
De Ribonuclease III [EC 3.1.26.3] (RNase III).
RM NC OR RNCS OR BH2489.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales/Bacilli group; Bacillus.
CX Bacillus/Staphylococcus group; Bacillus.
ON NCB_I_TaxId=8665;
RX SEQUENCE FROM N.A.
STRAIN=C-125 / JCM 9153;
MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kohara S.,
Horikoshi K.;
FT "Complete genome sequence of the alkaliphilic bacterium Bacillus
halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -FUNCTION: Digests double-stranded RNA. Involved in the processing
of ribosomal RNA precursors and of some mRNAs (by similarity).
CC -CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-phosphomonester.
CC -SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -SIMILARITY: CONTAINS 1 DBRH (DOUBLE-STANDED RNA-BINDING) DOMAIN.
CC -SIMILARITY: CONTAINS 1 RNASE III DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/isb-sib.ch).

--
EMBL: AP001515; BAB06208.1; -
InterPro: IPRO001159; DS_RBD.
DR InterPro: PRO000999; RNase_3.
DR Pfam: PF00035; dsrm; 1.
DR Pfam: PR00636; Ribonuclease_3; 1.
DR SMART: SM00358; DSRM; 1.
DR SMART: SM00535; RIIBOC; 1.
DR POSSITE: PS00137; DS_RBP; 1.
DR POSITE: PS00517; RNase_3_1; 1.
DR POSITE: PS00142; RNase_3_2; 1.
KW Hydrolyase; Nuclease; Endonuclease; RNA-binding; Complete proteome.

FT DOMAIN 35 164
FT FTCD 190 259 DRBM III.
SQ SEQUENCE 263 AA; 30224 MW; 61BB8EIC557C5S485 CRC64;

Query Match 3.8%; Score 282; DB 1; Length 263;
Best Local Similarity 30.9%; Pred.No.1.le-07;
Matches 77; Conservative 50; Mismatches 110; Indels 12; Gaps 5,

1084 HPLOTPEPNDRPOLITSVIOKLTEEFERAIQVIFTTHVLARRAFALT RTGVGFNH --LTIG 1141

```

Db      15 HSERRRPP---KRLFLTAQOQMFDELLRLTNLTFANKKLLVQVATHTSSYNEHRIQSC 71
      1142 HQHREFLGDSIMQVATVEYLFHPDHHGHLLTLRSVLNNKQAVAEELGMOEVAI 1201
      Db      72 DNERLEFLGDAVLEAVSYLKYAEQMSGDMTKLRASIYCEPSLAQIAELHFGELVL 131
      Qy      1202 TNDKRRPVG-LRTKTLADLESFIALYTDKDLVEYHTFEMVCFEPRLKEPILNQMDND 1260
      Db      1332 LGKGEEMTKRRPALIADVFESFGALYLDQMDAVILFERTYTPKISEGAPSH-RMD 190
      Qy      1261 PPSQLQOCCCLLTREGKEPDIPLYKTLQTVGSHARTYTVAVYFGERIGCGKGPISQQA 1320
      Db      191 FKSQLOEF-----IORDNLGHIHYEIVQERGPANHREFVSEVLNNETLIGVGTGRSKKEA 245
      Qy      1321 EMGAMMDL 1329
      Db      246 EQHAAQAL 254

```

RESULT 7

```

EXTN_TOBAC      STANDARD:      PRT:      620 AA.
AC      P13983;
DT      01-JAN-1990 (Rel. 13, Created)
DT      01-JAN-1990 (Rel. 13, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Extensin precursor (Cell wall hydroxyproline-rich glycoprotein).
GN      HRCGN13.
OS      Nicotiana tabacum (Common tobacco).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC      Asteridae; eusterids I; Solanales; Solanaceae; Nicotiana.
OX      NCBI_TaxID=4097;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=CV. XANTHT; TISSUE=leaf;
RA      MEDLINE=90128263; PubMed=2612909;
RT      Keller B., Lamb C.J.;
RT      "Specific expression of a novel cell wall hydroxyproline-rich
RT      glycoprotein gene in lateral root initiation.";
RL      Genes Dev. 3:1639-1646(1989).
CC      -1- FUNCTION: HAS A SPECIALIZED STRUCTURAL FUNCTION, POSSIBLY IN
CC      THE MECHANICAL PENETRATION OF THE CORTEX AND EPIDERMIS OF THE
CC      MAIN ROOT.
CC      -1- SUBCELLULAR LOCATION: Extracellular matrix.
CC      -1- PTM: EXTENSINS CONTAIN A CHARACTERISTIC REPEAT OF THE PENTAPEPTIDE
CC      SER-PRO(4). THE PROLINE RESIDUE IS HYDROXYLATED AND THEN
CC      GLYCOSYLATED.

```

```

CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL, X13885; CAA32090.1; -.
CC      PIR, S06733; S06733.
CC      Repeat; Cell wall; Glycoprotein; Signal; Structural protein;
CC      Hydroxylation.
CC      KW      SIGNAL
CC      FT      CHAIN      1      20      POTENTIAL.
CC      FT      REPEAT      70      73      EXTENSIN.
CC      FT      REPEAT      148      151      H-A-P-P.
CC      FT      DOMAIN      229      242      2 X 7 AA TANDEM REPEATS OF T-H-R-H-A-P-P.
CC      FT      REPEAT      229      235      1.
CC      FT      REPEAT      236      242      2.
CC      FT      DOMAIN      205      620      CONTAINS THE SER-PRO(4) REPEATS.
CC      FT      DOMAIN      499      600      3 X APPROXIMATE TANDEM REPEATS.
CC      FT      SEQUENCE      620      AA:      65406      MW:      641DD2278AB28524      CRC64;

```

```

Query Match      3.6%; Score 271.5; DB 1; Length 620;
Best Local Similarity 30.0%; Pred. No. 1e-06;
Matches 77; Conservative 19; Mismatches 80; Indels 81; Gaps 13;
      Qy      29 PSADSPRPQMLRLHPQPPVQYQYBPSPASTTFS-----NSPANEFLPPRPDVPFP- 82
      Db      293 PPSPIYSP-----PPPAVSPSPPPPTPTFFSPPPAVSPSPPTYSPPPTYLPLPS 342
      Qy      83 -----PPMPSAGCP-----LPPCP-----IRPFPMHOM-----RBP 110
      Db      343 SPIYSPSPPPVSPSPPPSYSPPPPTYLPPPPSSPPPSSTSPPPPTYEOSPPPPAVSP 402
      Qy      111 FVPVPCPP-----MPPMPCPNPVPBG-APPQGT-----PFMMPP----- 149
      Db      403 LPAPPTYSPPPPPTYSPPPTTYAQPPPLPPTYSPPPPAVSPSPPPPTYSPPPPAY 462
      Qy      150 SMPHPPPPVMPQAVNYQYPPGSHNFPSPNSQNPNSSFLPSANSSSPHFLHP 209
      Db      463 AQPPPPPTYSPPPPAVSPSPPSPIYSPPP-----QVQPLPPTSP-----PPPRILHP 514
      Qy      210 -----YPLKAPSERRSP 222
      Db      515 PPHQPPRPPTTYGQPP 531

```

RESULT 8

```

RNC_BUCAL      STANDARD:      PRT:      226 AA.
AC      P57346;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Ribonuclease III (EC 3.1.26.3) (RNase III).
GN      RNC OR BU258.
OS      Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
OS      symbiotic bacterium).
OC      Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX      NCBI_TaxID=118099;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=TOKYO 1998;
RA      MEDLINE=20445173; PubMed=10993077;
RT      Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT      "Genome sequence of the endocellular bacterial symbiont of aphids
RT      Buchnera sp. APS.";
RL      Nature 407:81-86(2000).
CC      -1- FUNCTION: DIGESTS DOUBLE-STRANDED RNA. INVOLVED IN THE PROCESSING
CC      OF RIBOSOMAL RNA PRECURSORS AND OF SOME MRNAS (BY SIMILARITY).
CC      -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC      phosphomonoester.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -1- SIMILARITY: CONTAINS 1 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.
CC      -1- SIMILARITY: CONTAINS 1 RNASE III DOMAIN.

```

```

CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL, AP001118; BAB12968.1; -.
CC      Interpro: IPR001159; DS_RBD.
CC      Interpro: IPR000999; RNase_3.
CC      Pfam: PF00035; dsrm. 1.
CC      Pfam: PF00636; Ribonuclease_3; 1.
CC      SMART: SM00358; DSRM; 1.
CC      SMART: SM00535; RIBOC; 1.
CC      PROSITE: PS01317; DS_RBD; 1.
CC      PROSITE: PS00517; RNASE_3_1; 1.
CC      PROSITE: PS0142; RNASE_3_2; 1.
CC      DR      Hydrolase; Nuclease; Endonuclease; RNA-binding; Complete proteome.
CC      KW

```



```

Db 851 VSKGS-POGRRLONQMRATKFNVLTTVEYVTKRAVLAKIQWKYITDEGHMKNH 909
Qy 774 -----HFGIRPAQSTAYADPOYOKIMKSYKRLHLA-----NSPKV 811
Db 910 CKLTVNLTWH-IAPYRLLLTGTPLQNKLPRLMALNLLPSIKSCSTFQWNNAPAT 968
Qy 812 QTKOKLAQREBALOKIROKNKTMRREYVELSSQGFMTGIRSDVCOHAMLPVLTNHR 871
Db 969 TGEKVELNEEETIL-----IIRR-----LHKVLRPFLRLRLK 1000
Qy 872 ---HYQCLMHLDKLIGYTFQ-DRCLLOLQAMTHPSHHLNFGMNPDPHARSLNSNGIRQPY 927
Db 1001 KEVEHQ-----LPDKVEYIIRKQMSALQ-----1023
Qy 928 GDRKVHNMKRGKINTLINISRLGODPTPSRINNELEFLGDAVEFLTSVHLTYL 987
Db 1024 ---RLYKHMOSKV--LITGSEKQ-----HG-----1047
Qy 988 FSLTEGLATYRTAIVONQHLAKKLELDPEMLYAH-----GPDLC 1031
Db 1048 -----KGGAKALMTIYO-----LRKLCNHPFW-FQHIIEKYCDHTGSHGVVSGDLY 1094
Qy 1032 RESDLRHMANCFEALIGAVILEGSLERAKQFGRLFNDDDLKRWMLNPLHLQLOEP 1091
Db 1095 RVS-----GKFE-----LDRIL--PKLKA-----1112
Qy 1092 NTRDROLIETSPVLOKLTFFEEAIG-VIFTHVL-----LARAFLTRVGFNHLT 1139
Db 1113 -TNHNVLECCMTQCMITIEDLGMROFGYLRLOSTTAEDRGELLKRFNKK--GSDVFW 1159
Qy 1140 LGHNOEMFLGDSIMQVATEYELTHPPDHGHILTLIRSSLVNNRQAKVAEELGMOEY 1199
Db 1170 FLSTRAGGLGIN-LQADTVYIFDSQWNPQODLOAQDRAHRIQGRNEVRY-----1219
Qy 1200 ATTNKTRPVGLRKTLDLIESFIAA-----LYTDKLEVYHNMNCFPRLEF---1252
Db 1220 -----LRMTVNSVEERILALARKLNDEKVIQAGMDQKSTGSEKQFLQOT 1267
Qy 1253 ILNODMNPKSQLOQ--OCCLTLRTEGKEPDILPKYKLOTQVPSHARTYTAVYFKGERI 1309
Db 1268 ILHODNEEENEVPPDEMINMIAESEELEIFKRDABERKK-----EDEEI 1316
Qy 1310 GCGKPSIOQAMGAA--DALKYNPQMAHOKRITGRKRYO-----ELKEKRW 1357
Db 1317 HGRERLIDESLPWMLTKDDEVERFYH--QYEDDTILGSGRSORKEVDYDLSLTEREW 1374
Qy 1358 -----EREHOREPPTEDIKK 1374
Db 1375 LKAIDGAEFEDEEDDSKRRRK 1399

RESULT 15
RNC_SALTY STANDARD: PRT: 226 AA.
AC 056056;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ribonuclease III (EC 3.1.26.3) (RNase III).
GN RNC OR STM2581.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TS117;
RX MEDLINE=97295239; PubMed=9150881;
RA Anderson P.E., Matsunaga J., Simons E.L., Simons R.W.;
RT "Structure and regulation of the Salmonella typhimurium rnc-era-reco
RT operon.";
RL Biochimie 78:1025-1034(1996).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvanev E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -1- FUNCTION: DIGESTS DOUBLE-STRANDED RNA. INVOLVED IN THE PROCESSING
CC OF RIBOSOMAL RNA PRECURSORS AND OF SOME MRNAS.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphonucleoside.
CC -1- SUBUNIT: ORGANIZED INTO A STRUCTURE (PROCESSOME) CONTAINING A
CC NUMBER OF RNA-PROCESSING ENZYMES.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: CONTAINS 1 DBM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 RNASE III DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, UA8415; AAA92440.1; -
CC EMBL, AE008817; AAL21473.1; -.
CC StvGene; SG10695; rnc.
CC InterPro; IPR001159; DS_RBD.
CC InterPro; IPR000999; RNase_3.
CC Pfam; PF00035; dsrm; 1.
CC Pfam; PF00636; Ribonuclease_3; 1.
CC SMART; SM00358; DSRM; 1.
CC SMART; SM00355; RIBOC; 1.
CC PROSITE; PS00137; DS_RBD; 1.
CC PROSITE; PS00517; RNase_3_1; 1.
CC PROSITE; PS0142; RNase_3_2; 1.
CC K1 Hydrolyase; Nuclease; Endonuclease; RNA-binding; Complete proteome.
CC FT DOMAIN 6 128
CC FT DOMAIN 208 224
CC FT CONFLICT 57 58 YH -> S (IN REF. 1).
CC FT CONFLICT 73 74 AR -> DP (IN REF. 1).
CC FT CONFLICT 217 220 EOAL -> NSV (IN REF. 1).
CC FT CONFLICT 226 AA; 25505 MW; 5752C0113C0A055A CRC64;
CC SQ

```

```

Query Match 3.4%; Score 258; DB 1; Length 226;
Best Local Similarity 29.5%; Pred. No. 1,4e-06;
Matches 69; Conservative 47; Mismatches 102; Indels 16; Gaps 6;

```

```

Qy 1101 SPVLOKLTFFEEAIGVTFTHVRLARAFRLRTGYFHNHLTGHNQRMFLGDSIMQVATE 1160
Db 2 NPVY--INRLQKLGITFNQELLOQLTHRSASSK-----HNERLEFGDSLISLVIAN 54
Qy 1161 YLFTHPPDHGHILTLIRSSLVNNRTOAKVAEELGMOEYATINDKTRKRVGLTKT-LAD 1219
Db 55 ALYHFRPVDEGMSRRRATLVGNLTALAEARFDGECRLRGPGRLKSGGFRRESILAD 114
Qy 1220 LLESFIALYTDKDLDEVHTFMVAVCFPPRLKEITLNDQWMDPKRSQLOCCCLTLRTGKRP 1279
Db 115 TVEALIGGVFLDSNIQVDELINWTKTRIDEISPDQKQDPKTRIQE-----YLQGRHL 169
Qy 1280 DIPYKTLQTVGSHARTYTAVYFPG--BRICGKPSIOQEMGAADALEK 1331
Db 170 PLPSYLVQYRGAEHDEFTIHCQVSGLSBPV-VGIGSSRRKRAEQAAADQALK 222

```

Search completed: October 9, 2002, 18:24:01
Job time: 41 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 17:13:02 ; Search time 66 Seconds
(without alignments)
2312.354 Million cell updates/sec

Title: US-09-900-425a-2

Perfect score: 7500
Sequence: 1 MMGGTCHRMSPHFGRGCP...MRWERHQREPEDEEDIKK 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4071	54.3	769	22	AAB92635	Human protein sequ
2	2963	39.5	1337	22	ABB58539	Drosophila melano
3	1961	26.1	378	22	AAU20587	Human secreted pro
4	1961	26.1	378	22	AAU21744	Human secreted pro
5	1956	26.1	378	22	AAU20385	Novel human neopla
6	1320	17.6	263	22	AAU21580	Human secreted pro
7	1295.5	17.3	301	22	AAAB6381	Novel human neopla
8	1289	17.2	267	22	AAAB6379	Human breast cance
9	1122	15.0	277	22	AAAB6383	Human breast cance
10	496	6.6	115	20	AAV12224	Human 5' EST seque
11	496	6.6	115	21	AA00554	Human secreted pro

12	340	4.5	1151	22	ABB61598	Drosophila melano
13	311	4.1	560	22	ABG21040	Novel human diagno
14	303.5	4.0	406	22	ABG27250	Novel human diagno
15	277.5	3.7	232	20	AAW81354	S. pneumoniae rnc
16	277	3.7	1537	22	ABB66989	Drosophila melano
17	276.5	3.7	1278	22	AAAM39259	Human polypeptide
18	275.5	3.7	594	22	ABB61362	Drosophila melano
19	273	3.6	973	22	AAW41045	Human polypeptide
20	272.5	3.6	702	22	ABB63064	Drosophila melano
21	267	3.6	243	20	AAV09366	Staphylococcus aur
22	265.5	3.5	470	22	ABG21932	Novel human diagno
23	261	3.5	572	18	AAW31855	Mycobacterium tube
24	261	3.5	763	18	AAW31852	S. epidermidis ope
25	260.5	3.5	245	22	AAAG8282	Drosophila melano
26	259.5	3.5	1638	22	ABB61946	Novel human diagno
27	259	3.5	1217	22	ABG09876	Novel human diagno
28	258	3.4	598	22	ABG14000	Drosophila melano
29	255.5	3.4	707	22	ABB68526	Drosophila melano
30	254	3.4	2249	22	ABB60977	Drosophila melano
31	254	3.4	2249	22	AAE09769	Drosophila melano
32	253	3.4	1239	22	ABG09877	Novel human diagno
33	245	3.3	565	22	AAW78694	Human protein sequ
34	244.5	3.3	446	22	ABB70063	Drosophila melano
35	242	3.2	255	18	AAW20772	H. pylori cytoplas
36	240	3.2	615	21	AAAG4152	Arabidopsis thalia
37	240	3.2	679	21	AAAG4151	Arabidopsis thalia
38	240	3.2	691	21	AAAG4150	Sugar beet chitina
39	239.5	3.2	439	13	AAK28150	Neospora NC-p65 ve
40	239.5	3.2	691	22	ABB03068	Human expressed po
41	238	3.2	861	22	AAU23180	Novel human diagno
42	238	3.2	325	22	ABG21919	Novel human diagno
43	236.5	3.2	1030	22	ABG23699	Novel human diagno
44	236	3.1	299	12	AAK13995	Rape extensin. Br
45	235.5	3.1	299	12	AAK13995	

ALIGNMENTS

RESULT 1
ID AAB92635 standard; Protein; 769 AA.
XX
AC AAB92635;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:10949.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
XX
PR 27-AUG-1999; 99JP-0300253.
XX
PR 11-JAN-2000; 2000JP-0118776.
XX
PR 02-MAY-2000; 2000JP-0183767.
XX
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto U;
XX
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -

PS Claim 8; SEQ ID 10949; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
CC AAH95893 represent human amino acid sequences; and AAH16229 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 769 AA:

Query Match 54.3%; Score 4071; DB 22; Length 769;
Best Local Similarity 99.5%; Pred. No. 1.9e-280;
Matches 765; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY . 606 MFAHAPLNIPLCKVIRINIDYTHFIEEMPEMFCVAGLEFSLFRLDLELYDMLK 665
DB 1 MFAHAPLNIPLCKVIRINIDYTHFIEEMPEMFCVAGLEFSLFRLDLELYDMLK 60
QY 666 GLFEDSPCCRFHFMRFVFLPDGKEVLSMHQILLYLRCSKALVPEEETANMLQW 725
DB 61 GLFEDSPCCRFHFMRFVFLPDGKEVLSMHQILLYLRCSKALVPEEETANMLQW 120
QY 726 EELEMQKYAECKGMIVTNGTKPSSVRIDOLREQFNPDVITPEIIVHSGIRPAQSYA 785
DB 121 EELEMQKYAECKGMIVTNGTKPSSVRIDOLREQFNPDVITPEIIVHSGIRPAQSYA 180
QY 786 GPPYQYKLMKSVYKLRHLANSRKYKQDKOKLADREBALQIKRQKMMREYVVELSSQ 845
DB 181 GPPYQYKLMKSVYKLRHLANSRKYKQDKOKLADREBALQIKRQKMMREYVVELSSQ 240
QY 846 GPMKIGIRSDVCOHAMMLPVLTNHTIRYQCLMHLDKLIGYTFODRCLQLAMTPHSNHLN 905
DB 241 GPMKIGIRSDVCOHAMMLPVLTNHTIRYQCLMHLDKLIGYTFODRCLQLAMTPHSNHLN 300
QY 906 FGMNDPHANSLNSGCIROPKYGRKVVHMMHMKKGINTLNIMSRQLQDDPTPSRINHN 965
DB 301 FGMNDPHANSLNSGCIROPKYGRKVVHMMHMKKGINTLNIMSRQLQDDPTPSRINHN 360
QY 966 EELEFLGDAVVEFLSVNHYLFFPSLEGGGLATYRTATVQNOHMLAKKLELDPMFLYA 1025
DB 361 EELEFLGDAVVEFLSVNHYLFFPSLEGGGLATYRTATVQNOHMLAKKLELDPMFLYA 420
QY 1026 HGPDLCSRESDLRRHMANCFEALIGAVVLEGSLEAKQDFGRLLFNDPDLREVLNYPPLH 1085
DB 421 HGPDLCSRESDLRRHMANCFEALIGAVVLEGSLEAKQDFGRLLFNDPDLREVLNYPPLH 480
QY 1086 LQLQEPNTDROLIETSPVLQKLTFFEEALIGVITFHVRLRLAARFTLRVGFNNHLTGNOR 1145
DB 481 LQLQEPNTDROLIETSPVLQKLTFFEEALIGVITFHVRLRLAARFTLRVGFNNHLTGNOR 540
QY 1146 MEFLGDSIMQVATERYLTFHPDHHGHGHTLLRSSLVNNRQAKVAEELGMEYATITDK 1205
DB 1146 MEFLGDSIMQVATERYLTFHPDHHGHGHTLLRSSLVNNRQAKVAEELGMEYATITDK 1205

DB 541 MEFLGDSIMQVATERYLTFHPDHHGHGHTLLRSSLVNNRQAKVAEELGMEYATITDK 600
QY 1206 TKRPVGLRTKTLADLESFALALYTPDKLEVYHTFMNVCFFPRLKEFILNDMPNPKSOL 1265
DB 601 TKRPVGLRTKTLADLESFALALYTPDKLEVYHTFMNVCFFPRLKEFILNDMPNPKSOL 660
QY 1266 QQCCLTLRTEGKEKPPDIPLYKTLQTVGSPHARTYVAVYFKGERIGCGKPSIQAEWGNA 1325
DB 661 QQCCLTLRTEGKEKPPDIPLYKTLQTVGSPHARTYVAVYFKGERIGCGKPSIQAEWGNA 720
QY 1326 MDLLEKYNPPQAHOKRFRFGKRYROELKEMRWEREHOEPEDETDIK 1374
DB 721 MDLLEKYNPPQAHOKRFRFGKRYROELKEMRWEREHOEPEDETDIK 769

RESULT 2

ID ABB58539 standard; Protein; 1327 AA.

AC ABB58539;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 2409.

KW Drosophila; developmental biology; cell signalling; insecticide;
pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001MO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PMD, Myers EM;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL02642.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
genes from Drosophila and for elucidating cell signalling and cell-cell
interactions -

PS Disclosure; SEQ ID NO 2409; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
CC publication, but was obtained in electronic format directly from WIPO
CC at http://wipo.int/pub/published_pct_sequences.

XX Sequence 1327 AA:

Query Match 39.5%; Score 2963; DB 22; Length 1327;
Best Local Similarity 45.6%; Pred. No. 2e-201;
Matches 621; Conservative 213; Mismatches 327; Indels 200; Gaps 32;

QY 90 QGPLPCCIRPPFPHQMRHPPVPCFPMPMPCPN-NPPVPGAP-----P 137
DB 3 QPLPPPPVQA-----PPPPPPPEEDLSPGCVGVPHNYSNNSHSG 46

CC disorders include for example: immune/autoimmune diseases (e.g. HIV
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, and
CC angina and thrombosis), infections caused by bacteria, viruses and
CC fungi and ocular disorders (e.g. corneal infections). (I) and (II),
CC agonists, antagonists and antibodies can also be used to promote wound
CC healing, maintain organs before transplantation, and support cell culture
CC of primary tissues. AAU20342-AAU20666 represent human secreted protein
CC amino acid sequences, and related sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification but was obtained in electronic format directly from WIPO
CC at: ftp.wipo.int/pub/publ/published_pct_sequences.
CC
XX
SO Sequence 378 AA;
Query Match 26.1%; Score 1961; DB 22; Length 378;
Best Local Similarity 98.9%; Pred. No. 5.3e-131;
Matches 374; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 997 ATYRAIVQNHLMALAKKLELDPMUYAHGPDLCRESDLRHANANCFEALIGAVYEGS 1056
Db 1 ATYRAIVQNHLMALAKKLELDPMUYAHGPDLCRESDLRHANANCFEALIGAVYEGS 60
Qy 1057 LEEAKQLFGRLLFNDPDLREVMYLPPLHQLQEPNTDROLIETSPYLQKITEPEAIGV 1116
Db 61 LEEAKQLFGRLLFNDPDLREVMYLPPLHQLQEPNTDROLIETSPYLQKITEPEAIGV 120
Qy 1117 IFTHVRLARAFYLTATVGFNHLITGHNRMEFLGDSIMQVATERYLTHPDDHGHSLTL 1176
Db 121 IFTHVRLARAFYLTATVGFNHLITGHNRMEFLGDSIMQVATERYLTHPDDHGHSLTL 180
Qy 1177 LRSSLVNNRTAKVAEEIGMOEYAITNDKTRPGYLRKTADLESFIALYDKDLEY 1236
Db 181 LRSSLVNNRTAKVAEEIGMOEYAITNDKTRPGYLRKTADLESFIALYDKDLEY 240
Qy 1237 VHTFMNVCFFPRLKEFILNODMNDPKSLOQCCLTLRTEGKEPDIPLYKTLQTVGPPSHAR 1296
Db 241 VHTFMNVCFFPRLKEFILNODMNDPKSLOQCCLTLRTEGKEPDIPLYKTLQTVGPPSHAR 300
Qy 1297 TYTVAVYFKGERIGCGKSPSTIOQAEMGAMDALEKYNPOMAHOKRFTGRKYRDELKEMR 1356
Db 301 TYTVAVYFKGERIGCGKSPSTIOQAEMGAMDALEKYNPOMAHOKRFTGRKYRDELKEMR 360
Qy 1357 WEREHOREPDETDIKK 1374
Db 361 WEREHOREPDETDIKK 378
RESULT 4
AAU21744
ID AAU21744 standard; Protein: 378 AA.
XX
AC AAU21744;
XX
DT 04-DEC-2001 (first entry)
XX
DE Novel human neoplastic disease associated polypeptide #177.
XX
KW Human; neoplastic disease associated polypeptide; cancer;
KW hyperproliferative disorder; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW neuroprotective; cytostatic; anti inflammatory; vasotropic.
XX
OS Homo sapiens.
XX
PN WO200155163-A1.
XX
PD 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01358.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227709.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465558/50.
DR N-PSDB; AAS34943.

XX Novel polypeptides and polynucleotides useful as diagnostic reagents to
PI diagnose diseases or disorders associated with aberrant expression or
PT activity of polypeptides, and for treating cancers, Rheumatoid
PT arthritis
XX
XX
PS Claim 11; SEQ ID NO 471; 687pp; English.
XX
CC The present invention relates to the isolation of novel human neoplastic
CC disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA
CC sequences encoding for these polypeptides. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of disorders involving neoplastic disease such as
CC hyperproliferative disorders (e.g. leukemia, bone cancer, bladder
CC cancer, brain stem glioma, adult liver cancer, childhood cerebellar
CC astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may
CC also be useful for treating other disorders such as neural disorders,
CC immune system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC and renal disorders. The polynucleotide sequences of the invention are
CC also useful in gene therapy. AAU21568-AAU21851 represent the novel human
CC neoplastic disease associated polypeptides of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 378 AA:
Query Match 26.1%; Score 1961; DB 22; Length 378;
Best Local Similarity 98.9%; Pred. No. 5.3e-11;
Matches 374; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 997 ATYRTAIYQNOHMLAKKLEIDPMLYAHGPDLCSRDLEHMANCFEALIGAVYLEGS 1056
DB 1 ATYRTAIYQNOHMLAKKLEIDPMLYAHGPDLCSRDLEHMANCFEALIGAVYLEGS 60
QY 1057 LEEAKOLGRLLENDPDLREYWLNPPLPLOJOPENTROLIETSPYOKTEFEBAIGV 1116
DB 61 LEEAKOLGRLLENDPDLREYWLNPPLPLOJOPENTROLIETSPYOKTEFEBAIGV 120
QY 1117 IFTHVRLARAFETLRTVGFNHLITGHNOMRMEFLDSIMQVATERYLFHFPDHHEGHLTL 1176
DB 121 IFTHVRLARAFETLRTVGFNHLITGHNOMRMEFLDSIMQVATERYLFHFPDHHEGHLTL 180
QY 1177 LRSSLVNNRTOAKVAEEIGMOEYATINDKTRPGLKTKTLADLESFIAALYTDKOLEY 1236
DB 181 LRSSLVNNRTOAKVAEEIGMOEYATINDKTRPGLKTKTLADLESFIAALYTDKOLEY 240
QY 1237 VHTFNNVCFPPRLKEFTLNOMNPNKSOLOCCJTLRTGKEPPIPLYTKLOTGSPSHAR 1296
DB 241 VHTFNNVCFPPRLKEFTLNOMNPNKSOLOCCJTLRTGKEPPIPLYTKLOTGSPSHAR 300
QY 1297 TYTVAVYFVGKGRIGCGKPSIOQAEMGAAMDALKEKYNPQAHOKRTFGKRYROELKEMR 1356
DB 301 TYTVAVYFVGKGRIGCGKPSIOQAEMGAAMDALKEKYNPQAHOKRTFGKRYROELKEMR 360
QY 1357 WEREHOEREPDETEDIK 1374
DB 361 WEREHOEREPDETEDIK 378
RESULT 5
AAU20385
ID AAU20385 standard; Protein; 378 AA.
XX
XX AAU20385;
XX
XX 06-DEC-2001 (first entry)
XX
XX Human secreted protein. Seq ID No 377.
DE
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
KW Rheumatoid arthritis; antiarteriosclerotic; cardiac; vascular;

PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0234400.
PR 14-SEP-2000; 2000US-0234401.
PR 14-SEP-2000; 2000US-0234406.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
FI
XX
XX WPI: 2001-46558/50.
DR N-PSDB; AAS34779.
XX
XX
XX Novel polypeptides and polynucleotides useful as diagnostic reagents to
PT diagnose diseases or disorders associated with aberrant expression or
PT activity of polypeptides, and for treating cancers, rheumatoid
PT arthritis -
XX
XX
PS Claim 11; SEQ ID No 307; 687pp; English.
XX
XX The present invention relates to the isolation of novel human neoplastic
CC disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA
CC sequences encoding for these polypeptides. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of disorders involving neoplastic disease such as
CC hyperproliferative disorders (e.g. leukaemia, bone cancer, bladder
CC cancer, brain stem glioma, adult liver cancer, childhood cerebellar
CC astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may
CC also be useful for treating other disorders such as neural disorders,
CC immune system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders
CC and renal disorders. The polynucleotide sequences of the invention are
CC also useful in gene therapy. AAU21568-AAU21851 represent the novel human
CC neoplastic disease associated polypeptides of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 263 AA;

Query Match 17.6%; Score 1320; DB 22; Length 263;

[illegible]

XX	RESULT 7
XX	AAB63281
XX	ID AAB63281 standard; Protein; 301 AA.
XX	AC AAB63281;
XX	DT 26-MAR-2001 (first entry)
XX	DE Human breast cancer associated antigen protein sequence SEQ ID NO:643.
XX	OS Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
XX	KW cancer associated antigen; cytostatic; cancer vaccine.
XX	MO Homo sapiens.
XX	PN MO200073801-A2.
XX	PD 07-DEC-2000.
XX	PE 26-MAY-2000; 2000MO-US14749.
XX	PR 28-MAY-1999; 9905-0136526.
XX	PR 10-SEP-1999; 9905-0153454.
XX	PA (LUDW-) LUDWIG INST CANCER RES.
XX	P1 Obata Y;
XX	DR WPI; 2001-025274/03.
XX	PT Nucleic acids encoding breast, gastric and prostate cancer associated
XX	PT antigen precursors, useful for diagnosing and treating a condition
XX	PT characterized by expression of an abnormal amount of a protein, e.g.
XX	PT cancer -
XX	PS Example 1; Page 499; 799pp; English.
XX	AA AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
XX	CC represent nucleotide sequences encoding human breast, gastric and
XX	CC prostate cancer associated antigen precursors (CAAP) respectively.
XX	CC AAB63332 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
XX	CC represent human breast, gastric and prostate CAAP protein sequence
XX	CC respectively. CAAPs have cytostatic activity and can be used in the
XX	CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
XX	CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
XX	CC condition characterised by expression of an abnormal amount of a protein,
XX	CC e.g. cancer.
XX	Sequence 301 AA;

Query Match	17.3%	Score 1295.5;	DB 22;	Length 301;
Best Local Similarity	89.7%;	Pred. No. 7.2e-84;		
Matches 252;	Conservative 4;	MissMatches 18;	Indels 7;	Gaps 4;
QY 860	AMMLPVLTHHHYHOCIMLMDLKIGTFDDROCLLOAMHPSHHLNFCGNPPHARNSLSN	919		
DB 15	AMMLPVLTHHHYHOCIMLMDLKIGTFDDROCLLOAMHPSHHLNFCGNPPHARNSLSN	74		
QY 920	CGIRPKYDRRKVVHHHHMKKGINTLINISMSLGGDDPPSRINNNERLEFLGDAVVEFL	979		
DB 75	CGIRPKYDRRKVVHHHHMKKGINTLINISMSLGGDDPPSRINNNERLEFLGDAVVEFL	134		
QY 980	TSVHLYYLFPSLEEGSLATYRTAIVONQMLAMLAKKLEDPMLYAHGPDLCRESDLRHA	1039		
DB 135	TSVHLYYLFPSLEEGSLATYRTAIVONQMLAMLAKKLEDPMLYAHGPDLCRESDLRHA	194		
QY 1040	MANCFEALIGAVYLBGSLEAKOLFGRLLFNPDLEEVYLANPLHPLQLOEPNTROLIG	1098		
DB 195	MANCFEALIGAVYLBGSLEAKOLFGRLLFNPDLEEVYLANPLHPLQLOEPNTROLIG	254		
QY 1099	ELSPVLQKLTFFE-ELIGVIFTHVRLAALFLRTYGFNHL	1138		
DB 255	NEQFYKKLTFFERNQGVNFL--LIGDFX-AAGGXHL	290		

RESULT 8
AAB63379
ID AAB63379 standard; Protein; 267 AA.
XX
AC AAB63379;
XX
DT 26-MAR-2001 (first entry)
XX
DE Human breast cancer associated antigen protein sequence SEQ ID NO:741.
XX
KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX
OS Homo sapiens.
OS
PN WO200073801-A2.
PN
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US14749.
XX
PR 28-MAY-1999; 99US-0136526.
PR 10-SEP-1999; 99US-0153454.
XX
PA (LUDW-) LUDWIG INST CANCER RES.
XX
PI
Obata Y;
XX
DR WPI; 2001-025274/03.
XX
PT Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer -
XX
XX
PS Example 1; Page 543; 799pp; English.
PS
XX
AAAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP) respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer.

XX Sequence 267 AA;

Query Match 17.2%; Score 1289; DB 22; Length 267;

Best Local Similarity 94.6%; Pred. No. 1.8e-83;

Matches 246; Conservative 2; Mismatches 10; Indels 2; Gaps 2;

QY 861 MMPLVLTNHRHQCLMHLDKLGYTFODRCILQLAMTHPSHNLNGMPDHARNSLNC 920
DB 1 MMPLVLTNHRHQCLMHLDKLGYTFODRCILQLAMTHPSHNLNGMPDHARNSLNC 60
QY 921 GIRQPKYGRKRVHNMRRKGINTLINIMSRIGQDDPTPSRINHNRELFGLDAVEEFLT 980
DB 61 GIRQPKYGRKRVHNMRRKGINTLINIMSRIGQDDPTPSRINHNRELFGLDAVEEFLT 120
QY 981 SVHLTYLFPSEGLATATATVQNOHLAMAKKLELDPFMYAHGPDLCSRESDLRHAM 1040
DB 121 SVHLTYLFPSEGLATATVQNOHLAMAKKLELDPFMYAHGPDLCSRESDLRHAM 180
QY 1041 ANCFEALIGAVYLEGSLEAKOLFGRLLFNDPDLREVWLNYPRLPQLQDEPNTDROLI-E 1099
DB 181 ANCFEALIGAVYLEGSLEAKOLFGRLLFNDPDLREVWLNYPRLPQLQDEPNTDROLI-E 240
QY 1100 TSPVLQKLEFE-EAIGVIF 1118
DB 241 FFOFYKKLTFEERNQGVNF 260

RESULT 9

AAB63383

ID AAB63383 standard; Protein; 277 AA.

AC AAB63383;

DT 26-MAR-2001 (first entry)

DE Human breast cancer associated antigen protein sequence SEQ ID NO:745.

DE Human breast cancer; gastric cancer; prostate cancer; diagnosis;

KW cancer associated antigen; cytostatic; cancer vaccine.

OS Homo sapiens.

PN WO200073801-A2.

PD 07-DEC-2000.

PF 26-MAY-2000; 2000WO-US14749.

PR 28-MAY-1999; 99US-0136526.

PR 10-SEP-1999; 99US-0153454.

XX (LUDW-) LUDWIG INST CANCER RES.

PA Obata Y;

PI Obata Y;

DR WPI; 2001-025274/03.

PT Nucleic acids encoding breast, gastric and prostate cancer associated
PT antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
PT cancer -

PS Example 1; Page 545; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
CC represent nucleotide sequences encoding human breast, gastric and
CC prostate cancer associated antigen precursors (CAAP), respectively.
CC AAB63232 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
CC represent human breast, gastric and prostate CAAP protein sequence
CC respectively. CAAPs have cytostatic activity and can be used in the
CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
CC acids or anti-CAAP antibodies are useful for diagnosing and treating a

CC condition characterised by expression of an abnormal amount of a protein,
CC e.g. cancer.

XX Sequence 277 AA;

Query Match 15.0%; Score 1122; DB 22; Length 277;

Best Local Similarity 83.1%; Pred. No. 1.4e-71;

Matches 231; Conservative 6; Mismatches 21; Indels 20; Gaps 4;

QY 942 INTLINIMSRIGQDDPTPSRINHNRELFGLDAVEEFLTSVHLTYLFPSEGLATYRT 1001
DB 1 INTLINIMSRIGQDDPTPSRINHNRELFGLDAVEEFLTSVHLTYLFPSEGLATYRT 60
QY 1002 AIVQNOHLAMAKKLELDPFMYAHGPDLCSRESDLRHAMANCFEALIGAVYLEGSLEAK 1061
DB 61 AIVQNOHLAMAKKLELDPFMYAHGPDLCSRESDLRHAMANCFEALIGAVYLEGSLEAK 120
QY 1062 QLFGRLLFNDPDLREVWLNYPRLPQLQDEPNTDROLIETSPVLQKLEFEBAIGVIFTHV 1121
DB 121 QLFGRLLFNDPDLREVWLNYPRLPQLQDEPNTDROLIETSPVLQKLEFEBAIGVIFTHV 180
QY 1122 RLARAFTLRVYGFVHNLTLGHNORMEFLGDSIMOLY-ATEYLFTHFPDHGGHTLLRSS 1180
DB 181 RLARAFTLRVYGFVHNLTLGHNORMEFLGDSIMOLYPOSTVFLS-QIIIGHLTFV--- 236
QY 1181 LVNNRTQAKVABELGMEYAITNDKTRKRPVGLRTYTLA 1218
DB 237 -----AKASLGE-----INKNFRPKGNKKRAMA 259

RESULT 10

AAV12224

ID AAV12224 standard; Protein; 115 AA.

AC AAV12224;

DT 18-JUN-1999 (first entry)

DE Human 5' EST secreted protein SEQ ID NO: 537.

KW Human; secreted protein; EST; expressed sequence tag; diagnosis;

KW forensic; gene therapy; chromosome mapping; signal peptide;

KW upstream regulatory sequence; cytokine activity; cell proliferation;

KW differentiation; haematopoiesis regulation; tissue growth regulation;

KW reproductive hormone regulation; chemokine; chemokine; haemostatic;

KW thrombolytic; anti-inflammatory; tumour inhibition.

OS Homo sapiens.

PN WO9906554-A2.

PD 11-FEB-1999.

PF 31-JUL-1998; 98WO-IB01238.

PR 01-AUG-1997; 97US-0905134.

XX (GEST) GENSET.

PA Ducleert A, Dumas Milne Edwards J, Lacroix B;

PI Ducleert A, Dumas Milne Edwards J, Lacroix B;

DR WPI; 1999-153784/13.

PS Claim 34; Page 599; 622pp; English.

XX New nucleic acids encoding human secreted proteins - obtained from
CC cDNA libraries prepared from kidney, fetal kidney, dystrophic
CC muscle, muscle and heart tissue
XX Claim 34; Page 599; 622pp; English.

CC	AAY11994	to AAY12260,	respectively.	The proteins given represent the		
CC	signal peptide	and an N-terminal fragment of a secreted protein.	The	nucleic acid sequences can be used for producing secreted human gene		
CC	products.	They can also be used to develop products for diagnosis and	therapy.	The proteins obtained may have cytokine activity, cell		
CC	proliferation/differentiation activity,	haematopoiesis regulating	activity,	tissue growth regulating activity,	reproductive hormone	
CC	regulating activity,	chemotactic/chemokinetic activity,	haemostatic and	thrombolytic activity,	receptor/ligand activity,	anti-inflammatory
CC	activity,	tumour inhibition activity or other activities.	The products	can be used in forensic,	gene therapy and chromosome mapping procedures.	CC
CC	The sequences can also be used for obtaining corresponding promoter	sequences.	The nucleic acids encoding the signal peptide can be used	for directing extracellular secretion of a polypeptide or the insertion	of a polypeptide into a membrane, or importing a polypeptide into	a cell.
SQ	Sequence	115 AA;				
Query Match		6.6%; Score 496; DB 20; Length 115;				
Best Local Similarity	95.7%; Pred. No. 1.2e-27;					
Matches	88; Conservative	2; Mismatches	2; Indels	0; Gaps		0
DQ	511 KKAHPDLHLEMYNPGOMNDGPCLCKSAKARRRGIGHSTYPEBEAIPCRPMTNAGR	570				
DQ	24 KRKLMDLHLEMYNPGOMNDGPCLCKSAKARRRGIRHSITYPEEAIKPCRPMTNAGR	83				
DQ	571 LFHYRIVSPPTNFPLDRPTVLEYDDHEIFE	602				
DQ	84 LFNHRIVSPPTNFPLDRPTVLEYDDHEIFE	115				
RESULT 11						
ID	AAG00554					
ID	AAG00554 standard; Protein; 115 AA.					
XX	AAG00554;					
DT	06-OCT-2000 (first entry)					
DE	Human secreted protein, SEQ ID NO: 4635.					
KW	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;					
KM	gene therapy; chromosome mapping.					
OS	Homo sapiens.					
PN	EP1033401-h2.					
XX	06-SEP-2000.					
PD	21-FEB-2000; 2000EP-0200610.					
PF	26-FEB-1999; 99US-0122487.					
PR	(GEST) GENSET.					
PA	Dumas Milne Edwards J, Duclert A, Giordano J;					
PI	WP1; 2000-500381/45.					
DR	N-PDB; AAC00560.					
XX	New nucleic acid that is a 5' expressed sequence tag (5' EST) for					
PT	obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for					
PT	diagnostic, forensic, gene therapy and chromosome mapping procedures -					
XX	Claim 13; SEQ ID 4635; 71pp + CD-ROM; English.					
CC	The present sequence is a polypeptide encoded by one of a large number					
CC	of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs					
CC	were prepared from total human RNAs or polyA+ RNAs derived from 30					
CC	different tissues. EST sequences usually correspond mainly to the 3'					
CC	untranslated region (UTR) of the mRNA because they are often obtained					

[illegible]

Db	6	TVLKNHFAIEADQKNLETAFTHTSVANHRLLKLSHNRLEEFADVAQOLLISLYLK	65
Qy	1166	FRDHESGHTLLNSSLVNNQTAQVAEEGMOEYATITNDKTRPVGLRKT--TLADLLES	1233
Db	66	YPKRPEBDLTKSLRAMIYREBSLGLAFARDQDOF--IKLGGEEKSGGRNPDITLIDAFEA	124
Qy	1224	FIALYIKDKOLEVYHTFMNCCFFPRAK--EFLINDMNDPKSOLDOCCLTIRTEKEDI	1281
Db	125	FLGLALLDCKVAAYKKEIYQVMIPKYEAEFEKIDTYD--KTHQDE--LQVNS--DV	175
Qy	1282	PL-FTKLQVQSHARTYTAAYVFKGERIGCGGSPITQOAEKGAMDALE	1331
Db	176	AIRQVVISFEGPAHDKFDEYVEVLBEKSTIGOGGNSKTLAEDEAKNAVEK	226

RESULT 2
 US-09-213-010-2
 Sequence 2, Application US/09213010
 Patent No. 6231630
 GENERAL INFORMATION:
 APPLICANT: Ionenlo, Michael A.
 APPLICANT: Rosenberg, Martin
 TITLE OF INVENTION: NOVEL COMPOUNDS
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Decheart Price & Rhoads
 STREET: 997 Lenox Drive, Building 3, Suite 210
 CITY: Lawrenceville
 STATE: NJ
 COUNTRY: USA
 ZIP: 08543
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSO for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/213,010
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/869,674
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Bloom, Allen
 REGISTRATION NUMBER: 29,135
 REFERENCE/DOCKET NUMBER: GM10013
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 609-520-3214
 TELEFAX: 609-520-3259
 TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 232 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-213-010-2

Query Match	3.7%;	Score 277.5;	DB 4;	Length 232;
Best Local Similarity	32.5%;	Pred. No. 3.6e-14;		
Matches	75;	Conservative 47;	Mismatches 92;	Indels 17;
				Gaps 8;

```

QY 1108 TEEEAIGVFTTHRIILAAFTLRTGPNF--TLGHNOMREIGSIOMQVATEYFLIH 1165
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 6 TVLNNHRAIEPADKNLLETFTFTSTANERRLKLSHNERLELPGAVLQILLISELYKK 65
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1166 PRDHGCHLTLFRRSLVNNRTQAKVAEELGEMQYATITNKTGRPVGLRKF--TLADLIES 1223
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 66 YPKRPEDLSLRLMIVREESLAFGRDQFOF--IKLGEGEKSSGRNRDIIIDAEVA 124
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1224 FIALYTDKDELYVHTGMVCFPRLLK--EFLINQDMDNPKSOLQOCCLLTFTEGKEDPI 1281
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```
D6      125 FLGALLDDKDAKKKEFLYQIMIKVNGEEMETTDY---KTHIQE---LLQYNG---DV 175
QY      1282 PL-IKTLOTVPSPHARTTYVAVFPEKGERIGCGCKPSTQAEAMGAADALEK 1331
          :   ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
D6      176 AIRQIVISETPAHMDKVDFVYLVEGSKIQQGGCRSKKLAEQAANAAVER 226
```

```

1      RESULT 3
2      US-08-965-087-2
3      : Sequence 2, Application US/08965087
4      : Patent No. 6346393
5      :
6      : GENERAL INFORMATION:
7      : APPLICANT: Fedon, Jason C.
8      : APPLICANT: Lonetto, Michael A.
9      : APPLICANT: Marra, Andrea
10     : APPLICANT: Palmer, Leslie M.
11     : APPLICANT: Rosenberg, Martin
12     : APPLICANT: Warren, Richard L.
13     : TITLE OF INVENTION: NOVEL RNCS
14     :
15     : NUMBER OF SEQUENCES: 4
16     :
17     : CORRESPONDENCE ADDRESS:
18     : ADDRESSEE: Dechert Price & Rhoads
19     : STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
20     : City: Philadelphia
21     :
22     : STATE: PA
23     :
24     : COUNTRY: US
25     :
26     : ZIP: 19103
27     :
28     : COMPUTER READABLE FORM:
29     : MEDIUM TYPE: Diskette
30     : COMPUTER: IBM Compatible
31     : OPERATING SYSTEM: DOS
32     :
33     : SOFTWARE: FASTSEQ for Windows Version 2.0
34     :
35     : CURRENT APPLICATION DATA:
36     : APPLICATION NUMBER: US/08/965,087
37     :
38     : FILING DATE:
39     : CLASSIFICATION: 435
40     :
41     : PRIOR APPLICATION DATA:
42     : APPLICATION NUMBER:
43     :
44     : FILING DATE:
45     : ATTORNEY/AGENT INFORMATION:
46     : NAME: Dickinson, Todd O
47     : REGISTRATION NUMBER: 28,354
48     : REFERENCE/DOCKET NUMBER: GM10116
49     : TELECOMMUNICATION INFORMATION:
50     : TELEPHONE: 215-994-2252
51     : TELEFAX: 215-994-2222
52     :
53     : TELEX:
54     : INFORMATION FOR SEQ ID NO: 2:
55     : SEQUENCE CHARACTERISTICS:
56     : LENGTH: 243 amino acids
57     : TYPE: amino acid
58     : STRANDEDNESS: single
59     : TOPOLOGY: linear
60     :
61     : US-08-965-087-2

```

Query Match	3.6%;	Score 267;	DB 4;	Length 243;
Best Local Similarity	30.5%;	Pred. No. 2.6e-13;		
Matches	75;	Conservative 55;	Mismatches 92;	Indels 24;
			Gaps	9;

```

QY 1092 NTBQLIETSVLOKLEFEBAIGVETVHLLARAFILRY--GNHLLGHNOBMEL 1149
Db 11 NRFRKRPD-----KME---LGYQONIDLQOAFSSHSIFINDNMNRDLHNERLEFL 61
QY 1150 GDSIMQVATPYELEIHPRDHHEGLILRLSSLVNNNTQAKVAELGMOEYATINDKTRP 1209
Db 62 GDVAVLELTVSRKYLEFKDRPNRPEGLNTRKMATICEPSLVFANKIGLNMELLG--KGEK 120
QY 1210 VGLRTK--TLADLLESTIALLYDKDLEYHTTMMNCCFPRLKEFLINDMW--DPSQL 1265
Db 121 TGRTRPSLLISDAEAFALIGALYDGDIDYWKAEKVIPEHNO--NELLGVNDKFT 177
QY 1266 QOCCILTRBEKEDIPLYTLDTQVPSHAKITYAVVYKGERIGCGRPSIQOAEEMA 1325

```



```

; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 905 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-357-014-9

Query Match      2.9%; Score 216; DB 4; Length 905;
Best Local Similarity 21.2%; Pred. No. 1.8e-08;
Matches 125; Conservative 58; Mismatches 212; Indels 196; Gaps 29;

QY 6 TCHMSFHP-----GRCPRGRGCGARSPARFRPQN-LRLHPQQRPVQYQYERP 56
DB 394 TCAALT-HRVPRLQPMGTCPTPARVPLLRHPSGPHRSILRACPCQMARPCQAPCP 452
QY 57 SA-----PSTFSNSPA-----PNFLPPR-----PDFVP 80
DB 453 SAGBPMAGVPSEPMWSTANLGLLSRPSVCPRLRPGENHRAGSNEDPILASGTP 512
QY 81 FPPMPPSAQ--GLRPPCIRPPFPNQMRRPP-----VPPCFPPMP 123
DB 513 -PPTIPDETFTGGVVP---RPAFVHYDKAEASDVEISLESDDSVIVYVEGGLPPLPP 567
QY 124 MCPNNRPVPGAPPGGCTPFPMMP-----PSMHPPPPVMPQOVNYQY 168
DB 568 PPSGATPP-PIAPTGP---PTASPVYPAKEEBELPAABGRLPPPPPPVPGVY---- 619
QY 169 PGYSHHNPSPSFSFNPPSSFLPSA-----NNSSPHFHRLPPYLPKAPSERR 220
DB 620 -----XLRPPQLVP-EGTPGGGPPALBEDLTVININSSDE-----EEEE 658
QY 721 SPERLKHYDDHNRHDSHGRGSRHSLDRERGRS---PDRRQDSRYSDYDRGRTSR 277
DB 659 EEEEEEEEEEEEEEEEEEDFEEDEDEEYFEEEEEEEEEEFE----- 711
QY 278 HRSYERERERERHRNRHRSRPSLSRKYKRGSGSYGLSVVPAGCTPELPGEI 337
DB 712 -----EEEGELE-EEEEDEEEEELEVEYEDLEFGTAG-----GEVEEGAPPPPTLPAL 761
QY 338 IKNTDSWAPPLEIVNHRSPSRKKRARWEERKDRMSDNQSSGDKNTYSIKKEPEETMP 397
DB 762 -----PPESPVPVQEPPEPGLLVEEPGTE-EEEGADTAPTL-----APEALP 807
QY 398 DKNEEEELLPVWIRCHSENYSDDPMQVGDSTVGTSLRLDYDK-----FE 449
DB 808 SOGEVERE-----GESPAAGPPP-QELVEEPPSXPTLLE 841
QY 450 EELSGROEKAARPPWEPKTKLDELSSESSEC--ESDEDSGSSSSD 498
DB 842 EETEDGSDKYQ-----PPETPAEEMETETEAALQEKEDDTAAMLAD 886

RESULT 11
US-08-574-959A-7
; Sequence 7, Application US/08574959A
; Patent No. 5962224
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Jung, Ratna K. Vadlamudi
; TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
; TITLE OF INVENTION: AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,959A
; FILING DATE: 19-DEC-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1135 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-574-959A-7

Query Match      2.9%; Score 216; DB 2; Length 1135;
Best Local Similarity 21.2%; Pred. No. 2.5e-08;
Matches 125; Conservative 58; Mismatches 212; Indels 196; Gaps 29;

QY 6 TCHMSFHP-----GRCPRGRGCGARSPARFRPQN-LRLHPQQRPVQYQYERP 56
DB 624 TCAALT-HRVPRLQPMGTCPTPARVPLLRHPSGPHRSILRACPCQMARPCQAPCP 682
QY 57 SA-----PSTFSNSPA-----PNFLPPR-----PDFVP 80
DB 683 SAGBPMAGVPSEPMWSTANLGLLSRPSVCPRLRPGENHRAGSNEDPILASGTP 742
QY 81 FPPMPPSAQ--GLRPPCIRPPFPNQMRRPP-----VPPCFPPMP 123
DB 743 -PPTIPDETFTGGVVP---RPAFVHYDKAEASDVEISLESDDSVIVYVEGGLPPLPP 797
QY 124 MCPNNRPVPGAPPGGCTPFPMMP-----PSMHPPPPVMPQOVNYQY 168
DB 798 PPSGATPP-PIAPTGP---PTASPVYPAKEEBELPAABGRLPPPPPPVPGVY---- 849
QY 169 PGYSHHNPSPSFSFNPPSSFLPSA-----NNSSPHFHRLPPYLPKAPSERR 220
DB 850 -----XLRPPQLVP-EGTPGGGPPALBEDLTVININSSDE-----EEEE 888
QY 221 SPERLKHYDDHNRHDSHGRGSRHSLDRERGRS---PDRRQDSRYSDYDRGRTSR 277
DB 889 EEEEEEEEEEEEEEEEEEDFEEDEDEEYFEEEEEEEEEEFE----- 941
QY 278 HRSYERERERERHRNRHRSRPSLSRKYKRGSGSYGLSVVPAGCTPELPGEI 337
DB 942 -----EEEGELE-EEEEDEEEEELEVEYEDLEFGTAG-----GEVEEGAPPPPTLPAL 991
QY 338 IKNTDSWAPPLEIVNHRSPSRKKRARWEERKDRMSDNQSSGDKNTYSIKKEPEETMP 397
DB 992 -----PPESPVPVQEPPEPGLLVEEPGTE-EEEGADTAPTL-----APEALP 1037
QY 398 DKNEEEELLPVWIRCHSENYSDDPMQVGDSTVGTSLRLDYDK-----FE 449
DB 1038 SOGEVERE-----GESPAAGPPP-QELVEEPPSXPTLLE 1071
QY 450 EELSGROEKAARPPWEPKTKLDELSSESSEC--ESDEDSGSSSSD 498
DB 1072 EETEDGSDKYQ-----PPETPAEEMETETEAALQEKEDDTAAMLAD 1116

```



```

RESULT 12
US-09-357-014-7
; Sequence 7, Application US/09357014
; Patent No. 6291645
; GENERAL INFORMATION:
; APPLICANT: Jaekyoon Shin, Insil Joung, Ratna K. Vadlamudi
; and Jack L. Strominger
; TITLE OF INVENTION: p62 POLYPEPTIDES, RELATED POLYPEPTIDES
; AND USES THEREFOR
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; City: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/357,014
; FILING DATE: 19-Jul-1999
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/574,959
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 7:
; LENGTH: 1135 amino acids
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-357-014-7

```

```

Query Match          2.9%; Score 216; DB 4; Length 1135;
Best Local Similarity 21.2%; Pred. No. 2.5e-08;
Matches 125; Conservative 58; Mismatches 212; Indels 196; Gaps 29;

```

```

QY 6 TCHRMSEFHP-----GRGCPRGSGHGAPASAFSPRON-LRLHPQOPVQOYQYEP 56
  || : || || || || || || || || || || || || || || || || || ||
DB 624 TCAALT-HPRVPLQMPGPTCPTRAPVPLRLPHRSPGPHSILKAPCPQMAPCQAPCP 682
QY 57 SA-----PSTFNSNPA-----DNFLPPR-----PDEVP 80
  || : || || || || || || || || || || || || || || || || ||
DB 683 SAGMPASGAPVPSPEPTSTANLIGLSRPSVCPRLLPENHRAGSNEDPILAPSGTP 742
QY 81 FPPMPPSAQ--GPLPCPIRPPRNQMHPP-----VPPCPPPPP 123
  || : || || || || || || || || || || || || || || || || ||
DB 743 -PPTLPDEFEGGRV--RPAVHYDKEDASDVETLSDDSDSVYIVPEGLPLPP 797
QY 124 MPCNNPPVPGAPGQGTFFPMMP-----PSMHPPPPPVMPQOVVYQY 168
  || : || || || || || || || || || || || || || || || || ||
DB 798 PPSGATFP-PIATIGP---FTASPPVPAKEPELPAAPGPLPPPPPPPPVGPV---- 849
QY 169 PPGTSHNPPPSFNSQNPPSFLPSA-----NNSSSPHFRLLPPYPLKAPBSER 220
  || : || || || || || || || || || || || || || || || || ||
DB 850 -----XLPPLQVLP-EGPFGGGGPPALEEDLVININSDE-----EEEE 888
QY 221 SPERLKHYDDHRRDHSHGGERHRSLSDRERGS---PDRRQDSSYRSQDYDGRKTPR 277
  || : || || || || || || || || || || || || || || || || ||
DB 889 EGEEEEEEEEEEEEEEEEDEEEEEDEBEYFEEEEEEEEEEEEFE----- 941

```

```

QY 278 HRSYRSRERERHRHRDRNRSPSLERSYKKKRGSGYGLSVYPACGTPELPGEI 337
  || : || || || || || || || || || || || || || || || || ||
DB 942 -----EEEGELE-EEEEDEDEEEELVEYEDLEFGAG-----GEVEBGAAPPPTLPAL 991
QY 338 IKNTDSWAPPLEIVNHRSPSREKKRARAWEEDKDRMSDQNSGKDKNYYSIKEREPEETMP 397
  || : || || || || || || || || || || || || || || || || ||
DB 992 -----PPESPAPVQPEPEPEPGLLVEEPEGTE-EEGADVAPTL-----APEALP 1037
QY 398 DKNEEEELKRVWIRCHSENYYSSDPMDQVGDSTVYGTSLRDYDK-----FE 449
  || : || || || || || || || || || || || || || || || || ||
DB 1038 SQGEVERE-----GESPAAGPP-DELVEEPSXPPTLLE 1071
QY 450 EELGSHQEKAKARPPPEPKTKLDELDSSSESEC--ESDSDSYCSSSD 498
  || : || || || || || || || || || || || || || || || || ||
DB 1072 EETEDSGDKVQ-----PPETPAEEMETPEALQKREDDPTAAMLAD 1116

```

```

RESULT 13
US-09-080-897-4
; Sequence 4, Application US/09080897
; Patent No. 5985574
; GENERAL INFORMATION:
; APPLICANT: King, Mary-Claire
; APPLICANT: Lynch, Eric D.
; APPLICANT: Lee, Ming
; APPLICANT: Morrow, Jan E.
; APPLICANT: Melosh, Piri L.
; APPLICANT: Leon, Pedro E.
; TITLE OF INVENTION: Modulators of Actin
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94010
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/080,897
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UW97-001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 343-4341
; TELEFAX: (650) 343-4342
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-080-897-4

```

```

Query Match          2.7%; Score 203; DB 2; Length 1255;
Best Local Similarity 32.1%; Pred. No. 3.1e-07;
Matches 63; Conservative 11; Mismatches 72; Indels 50; Gaps 10;

```

```

QY 58 ASTFSTNSPAPNFLPPR-----DFVFPFPPMPAPSAOGFLPCCPIRPPRNQMHPPPV 113
  || : || || || || || || || || || || || || || || || || ||
DB 562 APSVSSSAVPP--APRLGDSGTIVPPPPRPLRGVGVPPSPPLPGTCLPPPPPLDG 619
QY 114 PPGCFPP-----MPPMPCPN-----NPVFGAP-----PGCGTFFPMPPS 150
  || : || || || || || || || || || || || || || || || || ||
DB 620 GACIIPPQLPGSAAIAPPPLPPLGVASIPPPPLPGATAIAPPPLPGATAIP---PPPP 676

```


Search completed: October 9, 2002, 18:27:44
Job time : 46 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 18:24:08 : Search time 200 Seconds
(without alignments)
1900.652 Million cell updates/sec

Title: US-09-900-425A-2
Perfect score: 7500
Sequence: 1 MNGNTCHRMSEHPGRCPR.....MRWERHQRPEDETDIKK 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1006125 seqs, 276659714 residues

Total number of hits satisfying chosen parameters: 1006125

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/1/paa/PC1_NEW_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4071	54.3	769	5	US-09-629-469A-10949
2	2847	38.0	541	6	US-10-205-331-116
3	1961	26.1	378	5	US-09-764-882A-579
4	1961	26.1	378	6	US-10-103-313-471
5	1956	26.1	378	5	US-09-764-882A-377
6	1320	17.6	263	6	US-10-103-313-307
7	496	6.6	115	5	US-09-547-599C-3281
8	315	4.2	249	5	US-09-791-537-130915
9	309	4.1	266	5	US-09-791-537-130915
10	305	4.1	266	5	US-09-791-537-130915
11	286	3.8	480	6	US-10-216-583-1603
12	273	3.6	827	5	US-09-935-625-17077
13	273	3.6	847	5	US-09-935-625-17076
14	273	3.6	859	5	US-09-935-625-17076
15	273	3.6	879	5	US-09-935-625-17479
16	273	3.6	891	5	US-09-935-625-17075
17	273	3.6	903	5	US-09-935-625-17478
18	267	3.6	224	5	US-09-791-537-55409
19	265.5	3.5	1002	5	US-09-791-537-143780
20	263.5	3.5	1130	6	US-10-179-131-9979
21	262	3.5	226	7	US-09-791-537-19992
22	259.5	3.5	1589	5	US-09-935-625-9406
23	259.5	3.5	1589	5	US-09-935-625-9406
24	259.5	3.5	1589	5	US-09-935-625-9406
25	259.5	3.5	1589	5	US-09-935-625-9406
26	259.5	3.5	1615	5	US-09-935-625-26245

27	259.5	3.5	1649	5	US-09-935-625-9404	Sequence 9404, Ap
28	259.5	3.5	1649	5	US-09-935-625-26244	Sequence 26244, A
29	258	3.4	231	5	US-09-791-537-38286	Sequence 38286, A
30	258	3.4	598	1	PCT-US02-25485-2790	Sequence 2790, Ap
31	257	3.4	1250	1	PCT-US02-16876-9	Sequence 9, App11
32	256	3.4	246	5	US-09-791-537-130365	Sequence 130365,
33	254.5	3.4	648	6	US-10-179-131-5134	Sequence 5134, Ap
34	254	3.4	2249	5	US-09-866-557A-4	Sequence 4, App11
35	254	3.4	2249	5	US-09-858-862-4	Sequence 4, App11
36	254	3.4	2249	6	US-10-055-797-4	Sequence 130916,
37	251.5	3.4	239	5	US-09-791-537-130916	Sequence 80004, A
38	251.5	3.4	240	5	US-09-791-537-80004	Sequence 900, App
39	250.5	3.3	592	5	US-09-935-625-900	Sequence 5343, Ap
40	250.5	3.3	592	5	US-09-935-625-5343	Sequence 8099, Ap
41	250.5	3.3	592	5	US-09-935-625-8099	Sequence 25320, A
42	250.5	3.3	592	5	US-09-935-625-25320	Sequence 899, App
43	250.5	3.3	597	5	US-09-935-625-899	Sequence 5342, Ap
44	250.5	3.3	597	5	US-09-935-625-5342	Sequence 8098, Ap
45	250.5	3.3	597	5	US-09-935-625-8098	

ALIGNMENTS

RESULT 1
US-09-629-469A-10949
Sequence 10949, Application US/09629469A
GENERAL INFORMATION:
APPLICANT: OTA, TOSHIO
APPLICANT: ISOGAI, TAKAO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: HAYASHI, KOJI
APPLICANT: SATTO, KAORU
APPLICANT: YAMAMOTO, JUNICHI
APPLICANT: ISHII, SHIZUKO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: WAKAMATSU, AI
APPLICANT: NAGAI, KEIICHI
APPLICANT: OTSUKI, TETSUJI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
FILE REFERENCE: 084335/0123
CURRENT FILING DATE: 2000-07-28
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 1999-248036
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: JP 1999-300253
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 2000-241899
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/159, 590
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 60/183, 322
PRIOR FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 19025
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10949
LENGTH: 769
TYPE: PRT
ORGANISM: Homo sapiens
US-09-629-469A-10949

Query Match 54.3% Score 4071; DB 5; Length 769;
Best Local Similarity 99.5%; Pred. No. 4e-185;
Matches 765; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 606 MEAHPPLNIPICKRYIRNIDYTHFTFEMAMPENPCVKGLESLFLPRLDILEYDNK 665
DB 1 MEAHPPLNIPICKRYIRNIDYTHFTFEMAMPENPCVKGLESLFLPRLDILEYDNK 60

Oy	666	GPLEDSPCCRFHMFHMFVFEFLRDGGKEVLSHMOILLYLKSGKALVPEEELNMLOM	725
Db	61	GPLEDSPCCRFHMFHMFVFEFLRDGGKEVLSHMOILLYLKSGKALVPEEELNMLOM	120
Oy	726	BELEMOKTAEECKGKIIVTNPGRPSVVRIDOLDREOFNPDIITPFLIVHFGIRPOLSYA	785
Db	121	BELEMOKTAEECKGKIIVTNPGRPSVVRIDOLDREOFNPDIITPFLIVHFGIRPOLSYA	180
Oy	766	GDPOYOKLWKSIVYKLRHLLANSPKYQDQKOLAOREALOKIRQKNMREEVYELSSQ	845
Db	181	GDPOYOKLWKSIVYKLRHLLANSPKYQDQKOLAOREALOKIRQKNMREEVYELSSQ	240
Oy	846	GPWKIGISDVCQHMMPLVLTNHTIRYHNOCLNHLIDKLGSTYQDRCILQLOLMTHTSHHIN	905
Db	241	GPWKIGISDVCQHMMPLVLTNHTIRYHNOCLNHLIDKLGSTYQDRCILQLOLMTHTSHHIN	300
Oy	906	FGMNDHARNSISNCGIRPKYGDGKYNHMHNRKKGINTLITMSRLGDDPTPSKININ	965
Db	301	FGMNDHARNSISNCGIRPKYGDGKYNHMHNRKKGINTLITMSRLGDDPTPSKININ	360
Oy	966	ERLEFLGDAVVEFLTSVHLYLFLPSLEBGLATYRTAIVONOHMLAKKLELDPMLYA	1025
Db	361	ERLEFLGDAVVEFLTSVHLYLFLPSLEBGLATYRTAIVONOHMLAKKLELDPMLYA	420
Oy	1026	HGRDLCRESLDLHMANCFEALIGAVYLEGSLEBAKOLFGRLLFNDPDLREVWLNPLPH	1085
Db	421	HGRDLCRESLDLHMANCFEALIGAVYLEGSLEBAKOLFGRLLFNDPDLREVWLNPLPH	480
Oy	1086	LOLOEPNDROLIEMSPVLOKLTREEEAAGVFTFVRIILARAFTLRTVGFENHLLTGHONR	1145
Db	481	LOLOEPNDROLIEMSPVLOKLTREEEAAGVFTFVRIILARAFTLRTVGFENHLLTGHONR	540
Oy	1146	MEFLDSDIMQVATVATYLEFIHPDHEGHILTLRSSLVNNRTOAKVAEELGMOEYAITNDK	1205
Db	541	MEFLDSDIMQVATVATYLEFIHPDHEGHILTLRSSLVNNRTOAKVAEELGMOEYAITNDK	600
Oy	-1206	TKRPVGLRPTKTLADLLESTIALYTDKOLEYVHTFMNNOCFPRLEKFFLINOMNDPKSOL	1265
Db	601	TKRPVGLRPTKTLADLLESTIALYTDKOLEYVHTFMNNOCFPRLEKFFLINOMNDPKSOL	660
Oy	1266	OOCCLTLETKGEKPDIPLYKTLQOTVGPSSHARTYTAUVVYKGERIGCGKPSIOQAEEMGA	1325
Db	661	OOCCLTLETKGEKPDIPLYKTLQOTVGPSSHARTYTAUVVYKGERIGCGKPSIOQAEEMGA	720
Oy	1326	MDALEKYNPOMAHOKRFIGRKYRQELKMKMREHREHOEERPEDEDIKK 1374	
Db	721	MDALEKYNPOMAHOKRFIGRKYRQELKMKMREHREHOEERPEDEDIKK 769	
RESULT 2			
US-10-205-331-116			
; Sequence 116, Application US/10205331			
; GENERAL INFORMATION:			
; APPLICANT: Warner-Lambert Company			
; APPLICANT: Lee, Kevin			
; APPLICANT: Dixon, Alistair			
; APPLICANT: Brookspank, Robert			
; APPLICANT: Pinnock, Robert			
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain			
; FILE REFERENCE: WI-A-018199			
; CURRENT APPLICATION NUMBER: US/10/205,331			
; CURRENT FILING DATE: 2002-07-24			
; PRIOR APPLICATION NUMBER: GB 0118354.0			
; PRIOR FILING DATE: 2001-07-27			
; NUMBER OF SEQ ID NOS: 117			
; SEQ ID NO 116			
; SOFTWARE: PatentIn Ver. 2.1			
; LENGTH: 541			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
; FEATURE:			
; OTHER INFORMATION: Putative ribonuclease III			
US-10-205-331-116			

Query Match	38.0%	Score 2847	DB 6	Length 541
Best Local Similarity	99.3%	Pred. No. 2.8e-127		
Matches 537	Conservative	0	Mismatches 4	Indels 0
				Gaps 0
QY 834	MREYTVELSSGGFKTKGRSDVCSOHAMMLPVLTGHHIRYNOCIMLMDKLIGYTFODRCIL	893		
DB 1	MREYTVELSSGGFKTKGRSDVCSOHAMMLPVLTGHHIRYNOCIMLMDKLIGYTFODRCIL	60		
QY 894	QLAMTHPSHNLFGNNPDHARNSLSNCGIRQKYGDRKYNHNMKKGGINTLINTMSRLG	953		
DB 61	QLAMTHPSHNLFGNNPDHARNSLSNCGIRQKYGDRKYNHNMKKGGINTLINTMSRLG	120		
QY 954	ODDPFRRNNHNERLEFGDAVEEFLTSVNYLYLPPSLEEGGLATRYATVONOHILAMLA	1013		
DB 121	ODDPFRRNNHNERLEFGDAVEEFLTSVNYLYLPPSLEEGGLATRYATVONOHILAMLA	180		
QY 1014	KKLEIDPFMLYAHGPDLCRESDLRRHMANCFEALIGAVYLBGSLEBAKQLEGRLLFNDDP	1073		
DB 181	KKLEIDPFMLYAHGPDLCRESDLRRHMANCFEALIGAVYLBGSLEBAKQLEGRLLFNDDP	240		
QY 1074	LRWVNLNPLNHLQLOEPTDQOLIETSVYLOKLHFEFALIGVITTHVALLARATLRIV	1133		
DB 241	LRWVNLNPLNHLQLOEPTDQOLIETSVYLOKLHFEFALIGVITTHVALLARATLRIV	300		
QY 1134	GENHLLTIGNOMEFEGDSIMQVATETELFHPRDHNEHLLTLRSSLVNNTQAKVAEE	1193		
DB 301	GENHLLTIGNOMEFEGDSIMQVATETELFHPRDHNEHLLTLRSSLVNNTQAKVAEE	360		
QY 1194	IGMOEYATINDTKRPVGLRTKTLADLLESFTALAYTDKLEVYHTFNNVCFPRLEKEI	1253		
DB 361	IGMOEYATINDTKRPVGLRTKTLADLLESFTALAYTDKLEVYHTFNNVCFPRLEKEI	420		
QY 1254	LNQDNNDRPSLOQCCLTLTGTEGKRPDIPLXYTLQTVGSHARTVAVAYFGERIGCGK	1313		
DB 421	LNQDNNDRPSLOQCCLTLTGTEGKRPDIPLXYTLQTVGSHARTVAVAYFGERIGCGK	480		
QY 1314	GPSIOAEKGAMADLEKYNPOMAHOKRFTIGRKYQDLKEMRWERHOREBDETDIK	1373		
DB 481	GPSIOAEKGAMADLEKYNPOMAHOKRFTIGRKYQDLKEMRWERHOREBDETDIK	540		
QY 1374	K 1374			
DB 541	K 541			

PRIOR APPLICATION NUMBER: 60/225,757
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/226,868
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/216,647
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802

PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08

```

; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,065
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,398
```

```

Query Match          26.1%; Score 1961; DB 5; Length 378;
Best Local Similarity 98.9%; Pred. No. 1.8e-85;
Matches 374; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```

QY 997 ATYTAIVONQHMLAMAKKLEDPFMYLHAGPDLRESDLRHMANCFEALIGAVYLEGS 1056
    |||||
DB 1 ATYTAIVONQHMLAMAKKLEDPFMYLHAGPDLRESDLRHMANCFEALIGAVYLEGS 60
QY 1057 LEEAKOLFGRLLFNDPDLREVMNLPLHPLQLOEPNTDRLITSPVLQKTFEEAIGV 1116
    |||||
DB 61 LEEAKOLFGRLLFNDPDLREVMNLPLHPLQLOEPNTDRLITSPVLQKTFEEAIGV 120
QY 1117 IFTHVRLARAFTLRTVGFNHLTLGHNOEMFEGDSIMQVATEYLFHFPHHGHGHTL 1176
    |||||
DB 121 IFTHVRLARAFTLRTVGFNHLTLGHNOEMFEGDSIMQVATEYLFHFPHHGHGHTL 180
QY 1177 LRSSLVNNRTQAKVAEELGMOEYAITNDKTRPVGLRTKTLADLLESFIALYTDKLEY 1236
    |||||
DB 181 LRSSLVNNRTQAKVAEELGMOEYAITNDKTRPVGLRTKTLADLLESFIALYTDKLEY 240
QY 1237 VHTFMNVCFFPRLKEFLINQDMNDPKSLOQCCCLTRTEGKEPDIPLYKTLQTVGSPSHAR 1296
    |||||
DB 241 VHTFMNVCFFPRLKEFLINQDMNDPKSLOQCCCLTRTEGKEPDIPLYKTLQTVGSPSHAR 300
QY 1297 TYTVAVYFKGERIGCGKSPSIQOAEKGMADALEKYNPQMAHQKRFGRKYROELKEMR 1356
    |||||
DB 301 TYTVAVYFKGERIGCGKSPSIQOAEKGMADALEKYNPQMAHQKRFGRKYROELKEMR 360
QY 1357 WEREHOREPDEDEDIKK 1374
    |||||
DB 361 WEREHOREPDEDEDIKK 378
```

```

RESULT 4
US-10-103-313-471
; Sequence 471, Application US/10103313
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1207C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
```

```

; PRIOR Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 471
; LENGTH: 378
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-103-313-471
```

```

Query Match          26.1%; Score 1961; DB 6; Length 378;
Best Local Similarity 98.9%; Pred. No. 1.8e-85;
Matches 374; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```

QY 997 ATYTAIVONQHMLAMAKKLEDPFMYLHAGPDLRESDLRHMANCFEALIGAVYLEGS 1056
    |||||
DB 1 ATYTAIVONQHMLAMAKKLEDPFMYLHAGPDLRESDLRHMANCFEALIGAVYLEGS 60
QY 1057 LEEAKOLFGRLLFNDPDLREVMNLPLHPLQLOEPNTDRLITSPVLQKTFEEAIGV 1116
    |||||
DB 61 LEEAKOLFGRLLFNDPDLREVMNLPLHPLQLOEPNTDRLITSPVLQKTFEEAIGV 120
QY 1117 IFTHVRLARAFTLRTVGFNHLTLGHNOEMFEGDSIMQVATEYLFHFPHHGHGHTL 1176
    |||||
DB 121 IFTHVRLARAFTLRTVGFNHLTLGHNOEMFEGDSIMQVATEYLFHFPHHGHGHTL 180
QY 1177 LRSSLVNNRTQAKVAEELGMOEYAITNDKTRPVGLRTKTLADLLESFIALYTDKLEY 1236
    |||||
DB 181 LRSSLVNNRTQAKVAEELGMOEYAITNDKTRPVGLRTKTLADLLESFIALYTDKLEY 240
QY 1237 VHTFMNVCFFPRLKEFLINQDMNDPKSLOQCCCLTRTEGKEPDIPLYKTLQTVGSPSHAR 1296
    |||||
DB 241 VHTFMNVCFFPRLKEFLINQDMNDPKSLOQCCCLTRTEGKEPDIPLYKTLQTVGSPSHAR 300
QY 1297 TYTVAVYFKGERIGCGKSPSIQOAEKGMADALEKYNPQMAHQKRFGRKYROELKEMR 1356
    |||||
DB 301 TYTVAVYFKGERIGCGKSPSIQOAEKGMADALEKYNPQMAHQKRFGRKYROELKEMR 360
QY 1357 WEREHOREPDEDEDIKK 1374
    |||||
DB 361 WEREHOREPDEDEDIKK 378
```

```

RESULT 5
US-09-764-882a-377
; Sequence 377, Application US/09764882a
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT215
; CURRENT APPLICATION NUMBER: US/09/764,882a
; CURRENT FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
```


PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,267
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/216,880
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/225,270
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/251,869
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/235,834
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/234,274
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/234,223
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: 60/228,924
PRIOR FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/224,518
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,369
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/224,519
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,964
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/241,809
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/249,299
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/236,327
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/241,785
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/244,617
PRIOR FILING DATE: 2000-11-01
PRIOR APPLICATION NUMBER: 60/225,268
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/236,368
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/251,856
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/251,868
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/229,344
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/234,997
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: 60/229,343
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,345
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,287
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 60/229,513
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/231,413
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/229,509
PRIOR FILING DATE: 2000-09-05
PRIOR APPLICATION NUMBER: 60/236,367
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/237,039
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,038
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/236,370
PRIOR FILING DATE: 2000-09-29
PRIOR APPLICATION NUMBER: 60/236,802
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,037
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: 60/237,040
PRIOR FILING DATE: 2000-10-02

PRIOR APPLICATION NUMBER: 60/240,960
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/239,935
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/239,937
PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 60/241,787
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/246,474
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/246,532
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: 60/249,216
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,210
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/226,681
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,759
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/225,213
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/227,182
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 60/225,214
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/235,836
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/230,438
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/215,135
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 60/225,266
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/249,218
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,208
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,213
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,212
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,207
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,245
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,244
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,217
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,211
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,215
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,264
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,214
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/249,297
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/232,400
PRIOR FILING DATE: 2000-09-14
PRIOR APPLICATION NUMBER: 60/231,242
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,081
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/232,080
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,414
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/231,244
PRIOR FILING DATE: 2000-09-08
PRIOR APPLICATION NUMBER: 60/233,064

```

; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,065
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,398

```

```

Query Match      26.1%; Score 1956; DB 5; Length 378;
Best Local Similarity 98.7%; Pred. No. 3,1e-85;
Matches 373; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 997 ATYRTAIYONOHAMLAKELELDPEMLYAHGPDLCRESDLRIAMANCEPALLGANYLEGS 1056
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 1 ATYRTAIYONOHAMLAKELELDPEMLYAHGPDLCRESDLRIAMANCEPALLGANYLEGS 60
QY 1057 LEEAKOLFGRLLFNPDLEREVMNLNPLHPLQLQEPNTDROLIETSPVLOKTEFEBAIGV 1116
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 61 LEEAKOLFGRLLFNPDLEREVMNLNPLHPLQLQEPNTDROLIETSPVLOKTEFEBAIGV 120
QY 1117 ITHHRLARAFLLRTVGFNNHLLTGHNOEMELGDSIMQVATEYLFTHFPDHHGHLLTL 1176
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 121 ITHHRLARAFLLRTVGFNNHLLTGHNOEMELGDSIMQVATEYLFTHFPDHHGHLLTL 180
QY 1177 LKSSLVNNTQAKVAEELGMOEYAITNDKTRPVGLRRTKTLADLESFIALYTDKDEYV 1236
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 181 LKSSLVNNTQAKVAEELGMOEYAITNDKTRPVGLRRTKTLADLESFIALYTDKDEYV 240
QY 1237 VTFEPMNVCFPRPKKEFTLNQMDNPKSOLQOCCFLTRTEGKEPDIPLYKTIQTVGSPSHAR 1296
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 241 VTFEPMNVCFPRPKKEFTLNQMDNPKSOLQOCCFLTRTEGKEPDIPLYKTIQTVGSPSHAR 300
QY 1297 YTVAVYFEGERIGCGKSPSIQOAEEMGAMDALEKYNFPOMAHOKRFTGRKYOELKEMRW 1356
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 301 YTVAVYFEGERIGCGKSPSIQOAEEMGAMDALEKYNFPOMAHOKRFTGRKYOELKEMRW 360
QY 1357 WEREHOEREPDETEDIK 1374
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 361 WEREHOEREPDETEDIK 378

```

```

RESULT 6
US-10-103-313-307
; Sequence 307, Application US/10103313
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1207C1
; CURRENT APPLICATION NUMBER: US/10/103,313
; CURRENT FILING DATE: 2002-03-12
; NUMBER OF SEQ ID NOS: 653
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 307
; LENGTH: 263
; TYPE: PRT

```

```

; ORGANISM: Homo sapiens
US-10-103-313-307

```

```

Query Match      17.6%; Score 1320; DB 6; Length 263;
Best Local Similarity 97.7%; Pred. No. 2.9e-35;
Matches 251; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 1118 FTHHRLARAFLLRTVGFNNHLLTGHNOEMELGDSIMQVATEYLFTHFPDHHGHLLTL 1177
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 7 FTHHRLARAFLLRTVGFNNHLLTGHNOEMELGDSIMQVATEYLFTHFPDHHGHLLTL 66
QY 1178 RSSLVNNNTQAKVAEELGMOEYAITNDKTRPVGLRRTKTLADLESFIALYTDKDEYV 1237
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 67 RSSLVNNNTQAKVAEELGMOEYAITNDKTRPVGLRRTKTLADLESFIALYTDKDEYV 126
QY 1238 HTEPMNVCFPRPKKEFTLNQMDNPKSOLQOCCFLTRTEGKEPDIPLYKTIQTVGSPSHAR 1297
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 127 HTEPMNVCFPRPKKEFTLNQMDNPKSOLQOCCFLTRTEGKEPDIPLYKTIQTVGSPSHAR 186
QY 1298 YTVAVYFEGERIGCGKSPSIQOAEEMGAMDALEKYNFPOMAHOKRFTGRKYOELKEMRW 1357
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 187 YTVAVYFEGERIGCGKSPSIQOAEEMGAMDALEKYNFPOMAHOKRFTGRKYOELKEMRW 246
QY 1358 WEREHOEREPDETEDIK 1374
    ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
DB 247 WEREHOEREPDETEDIK 263

```

```

RESULT 7
US-09-547-599C-3281
; Sequence 3281, Application US/09547599C
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
; FILE REFERENCE: GEN-7119C1
; CURRENT APPLICATION NUMBER: US/09/547,599C
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: US 08/905,223
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,135
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,051
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,144
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,279
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/904,468
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,134
; PRIOR FILING DATE: 1997-08-01
; PRIOR APPLICATION NUMBER: US 08/905,133
; PRIOR FILING DATE: 1997-08-01
; NUMBER OF SEQ ID NOS: 3475
; SOFTWARE: Patent.pm
; SEQ ID NO 3281
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo Sapiens
; OTHER INFORMATION: kidney
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -23, -1
; OTHER INFORMATION: Von Heljne matrix
; OTHER INFORMATION: score 5.90
; FEATURE:
; OTHER INFORMATION: seq TTKFTLLQKSNA/KR
; NAME/KEY: UNSURE
; LOCATION: 3

```

OTHER INFORMATION: Xaa = Ala,Pro
US-09-547-599C-3281

Query Match 6.6%; Score 496; DB 5; Length 115;
Best Local Similarity 95.7%; Pred. No. 1.3e-16;
Matches 88; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 511 KKAHPRLHDELWYNDPGQNDGPKCSAKARTGIRHSIYGEBAIRKCPRPMTNAGR 570
DB 24 KXNLRDLHDELWYNDPGQNDGPKCSAKARTGIRHSIYGEBAIRKCPRPMTNAGR 83
QY 571 LFIHRTVSPPTNFDPRPVIEYDDHEIYFE 602
DB 84 LFHRTVSPPTNFDPRPVIEYDDHEIYFE 115

RESULT 8
US-09-791-537-130915

Sequence 130915, Application US/09791537

GENERAL INFORMATION:

APPLICANT: Blonomix, Inc.

APPLICANT: Debe, Derek

APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 130915

LENGTH: 249

TYPE: PRT

ORGANISM: Bacillus subtilis

US-09-791-537-130915

Query Match 4.2%; Score 315; DB 5; Length 249;
Best Local Similarity 34.7%; Pred. No. 1.2e-07;
Matches 85; Conservative 46; Mismatches 96; Indels 18; Gaps 7;

QY 1104 LOKLTFEEAIGVITTHVRLAARFTLTGVGFNHLTGH--NORMEFLGDSIMOLVATEY 1161
DB 17 VEOFKFEQERISVHFQNEKLLYQAFTHSSYVNEHRRKRPYEDNERLEFLGDVLELTISR 76
QY 1162 LFHFPRDHHEGLTLRSSLVNNRTQAKVAEELGMOEYAITNDKTRPVG-LRTKTLADL 1220
DB 77 LFPKYPAMSEGDITKLRAALVCEPSLVSLAHLSFGDLVLGKGEWMTGRRKRPALLADV 136
QY 1221 LESFIALYTDKDELYVTFMNVCFFPRLKEFLINQDNDPKSLOQCCLTLRTEGK--- 1277
DB 137 FFAFIALYLDQGLPEVSPFLKVVYFKINDGAPRHV-DRKSQLOE---YVQDQKGS 192
QY 1278 EFDIPLYKTLQTVGSPHARTYVAVYFKGERIGCGKPSIOAEMGAMDALEKYNF 1337
DB 193 E-----YKISNKGPAHNRFEALVSLKGEPLGVNGRSKKEAEOHAOEAALAKL---OK 244
QY 1338 AHOKR 1342
DB 245 HHTKQ 249

RESULT 9

US-09-791-537-62108

Sequence 62108, Application US/09791537

GENERAL INFORMATION:

APPLICANT: Blonomix, Inc.

APPLICANT: Debe, Derek

APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 62108
LENGTH: 266
TYPE: PRT
ORGANISM: Bacillus subtilis
US-09-791-537-62108

Query Match 4.1%; Score 309; DB 5; Length 266;
Best Local Similarity 34.6%; Pred. No. 2.5e-07;
Matches 84; Conservative 47; Mismatches 94; Indels 18; Gaps 7;

QY 1104 LOKLTFEEAIGVITTHVRLAARFTLTGVGFNHLTGH--NORMEFLGDSIMOLVATEY 1161
DB 17 VEOFKFEQERISVHFQNEKLLYQAFTHSSYVNEHRRKRPYEDNERLEFLGDVLELTISR 76
QY 1162 LFHFPRDHHEGLTLRSSLVNNRTQAKVAEELGMOEYAITNDKTRPVG-LRTKTLADL 1220
DB 77 LFPKYPAMSEGDITKLRAALVCEPSLVSLAHLSFGDLVLGKGEWMTGRRKRPALLADV 136
QY 1221 LESFIALYTDKDELYVTFMNVCFFPRLKEFLINQDNDPKSLOQCCLTLRTEGK--- 1277
DB 137 FFAFIALYLDQGLPEVSPFLKVVYFKINDGAPRHV-DRKSQLOE---YVQDQKGS 192
QY 1278 EFDIPLYKTLQTVGSPHARTYVAVYFKGERIGCGKPSIOAEMGAMDA---LEKYNF 1334
DB 193 E-----YKISNKGPAHNRFEALVSLKGEPLGVNGRSKKEAEOHAOEAALAKLKHHT 247
QY 1335 PQM 1337
DB 248 KQL 250

RESULT 10

US-09-791-537-104336

Sequence 104336, Application US/09791537

GENERAL INFORMATION:

APPLICANT: Blonomix, Inc.

APPLICANT: Debe, Derek

APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 104336

LENGTH: 266

TYPE: PRT

ORGANISM: Bacillus subtilis

US-09-791-537-104336

Query Match 4.1%; Score 305; DB 5; Length 266;
Best Local Similarity 34.2%; Pred. No. 3.8e-07;
Matches 83; Conservative 48; Mismatches 94; Indels 18; Gaps 7;

QY 1104 LOKLTFEEAIGVITTHVRLAARFTLTGVGFNHLTGH--NORMEFLGDSIMOLVATEY 1161
DB 17 VEOFKFEQERISVHFQNEKLLYQAFTHSSYVNEHRRKRPYEDNERLEFLGDVLELTISR 76
QY 1162 LFHFPRDHHEGLTLRSSLVNNRTQAKVAEELGMOEYAITNDKTRPVG-LRTKTLADL 1220
DB 77 LFPKYPAMSEGDITKLRAALVCEPSLVSLAHLSFGDLVLGKGEWMTGRRKRPALLADV 136
QY 1221 LESFIALYTDKDELYVTFMNVCFFPRLKEFLINQDNDPKSLOQCCLTLRTEGK--- 1277
DB 137 FFAFIALYLDQGLPEVSPFLKVVYFKINDGAPRHV-DRKSQLOE---YVQDQKGS 192
QY 1278 EFDIPLYKTLQTVGSPHARTYVAVYFKGERIGCGKPSIOAEMGAMDA---LEKYNF 1334
DB 193 E-----YKISNKGPAHNRFEALVSLKGEPLGVNGRSKKEAEOHAOEAALAKLKHHT 247


```

Db 643 SDKLSKRSVYRHHGSGMSVENSSEGRSPVSSKVKD-SEQVEKENSDLDANLSCDSKD 701
QY 455 ---ROEKAKAARPPWEPKTKLDEDESSSECESEDSTCCSSSDSEVFDVIAETIKRK 511
Db 702 TIRHQIKDKNRR-----KNK-----RSSREVSDD-----NGSSDSV-DDRKEAKRR 744
QY 512 KAHPRDLDELWYNDPGQMDGFLCKCSAKARRTGIRH 549
Db 745 RKEEKKTRKE-----EKRRRREERH 764

```

```

RESULT 13
US-09-935-625-17076
; Sequence 17076, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 17076
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..847
; OTHER INFORMATION: Ceres Seq. ID no. 2708207
US-09-935-625-17076

```

```

Query Match          3.6%; Score 273; DB 5; Length 847;
Best Local Similarity 24.8%; Pred. No. 4,6e-05;
Matches 158; Conservative 54; Mismatches 216; Indels 210; Gaps 32;

```

```

QY 8 HRMSFHRGCGCPRGCGHGAAPSAPSR-----PQNLRLHPQP-----PVQYQYE 54
Db 261 HRRPTHEGRGROSPAPSRRRRSPSPARRRRSPSPARRRRSPSPARRRRSPSPARRRR 320
QY 55 PRSAPSTFNSNAPNLFRLPRDFVPRP-----PMPSSAGCLPCCPI-----RPPRN 104
Db 321 SPSPARRHRSP-----PARRRRSPSPARRRRSPSP-ARRRRSPSLYRRNRSPSL 374
QY 105 HQM-----RHFPVPCFPMPRPMPCPN-----NPPV-----PGAP 137
Db 375 YRRNRSRSLAKRGSDSGRSPVARRLDPTGARLPSPSTEQRLPSPPVAQRLPSPP 434
QY 138 GCGTFFPMMPSPMPHPPPP-----VMPQVNTQYPPGYSH--HNFPPSFNSQNNPS 190
Db 435 RRAGLP--SPPAQRLPSPPRRAGLPSPMRI-----GGSNAANHLESPPS-----PS 479
QY 191 SFLPSANNSSPHFRHLPRPLPKAPSERKSPERLKHVDHHRHDSH-----G 239
Db 480 SLSPGGRK-----KVLSPVPVRRRSLTPDEERKSLSGGRTSPSHIKODGSMSPYRG 533
QY 240 RGERHRSLLDRERGRSPDRRQDSRYRSDYDGRTPSRHRSYERSERERERHRRHNDNR 299
Db 534 RKGSSPS-SRHQAKARSPVRRSP-----TPVNRRS-RRSSASARSDDRR--RR 578
QY 300 SPFLERSYKKEKKRSGRSTGLSVPEPAGCTPDLPGELITKNTDSMAPPLEIVNHRSPS-R 358
Db 579 SPSSSSPSR-----SRSP--VLHRSPPSR 602
QY 359 EKKRARWEEKDRMSDNO---SSGDKNTSTIKEKEPEETMDK-----NEEEEEE 406
Db 603 GRKHORERRSPGRSLSEODRVONSKLLKRTSVPTDKRKQLEKLLGVGRVETKQDERK 662
QY 407 LKRPWIRCTHSENYSSDPMDOVGSDTVVGTSLRLDYDKFEELGS-----454
Db 663 SDKLSKRSVYRHHGSGMSVENSSEGRSPVSSKVKD-SEQVEKENSDLDANLSCDSKD 721

```

```

QY 455 ---ROEKAKAARPPWEPKTKLDEDESSSECESEDSTCCSSSDSEVFDVIAETIKRK 511
Db 722 TIRHQIKDKNRR-----KNK-----RSSREVSDD-----NGSSDSV-DDRKEAKRR 764
QY 512 KAHPRDLDELWYNDPGQMDGFLCKCSAKARRTGIRH 549
Db 765 RKEEKKTRKE-----EKRRRREERH 784

```

```

RESULT 14
US-09-935-625-17480
; Sequence 17480, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 17480
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..859
; OTHER INFORMATION: Ceres Seq. ID no. 3023746
US-09-935-625-17480

```

```

Query Match          3.6%; Score 273; DB 5; Length 859;
Best Local Similarity 24.8%; Pred. No. 4,7e-05;
Matches 158; Conservative 54; Mismatches 216; Indels 210; Gaps 32;

```

```

QY 8 HRMSFHRGCGCPRGCGHGAAPSAPSR-----PQNLRLHPQP-----PVQYQYE 54
Db 273 HRRPTHEGRGROSPAPSRRRRSPSPARRRRSPSPARRRRSPSPARRRRSPSPARRRR 332
QY 55 PRSAPSTFNSNAPNLFRLPRDFVPRP-----PMPSSAGCLPCCPI-----RPPRN 104
Db 333 SPSPARRHRSP-----PARRRRSPSPARRRRSPSP-ARRRRSPSLYRRNRSPSL 386
QY 105 HQM-----RHFPVPCFPMPRPMPCPN-----NPPV-----PGAP 137
Db 387 YRRNRSRSLAKRGSDSGRSPVARRLDPTGARLPSPSTEQRLPSPPVAQRLPSPP 446
QY 138 GCGTFFPMMPSPMPHPPPP-----VMPQVNTQYPPGYSH--HNFPPSFNSQNNPS 190
Db 447 RRAGLP--SPPAQRLPSPPRRAGLPSPMRI-----GGSNAANHLESPPS-----PS 491
QY 191 SFLPSANNSSPHFRHLPRPLPKAPSERKSPERLKHVDHHRHDSH-----G 239
Db 492 SLSPGGRK-----KVLSPVPVRRRSLTPDEERKSLSGGRTSPSHIKODGSMSPYRG 545
QY 240 RGERHRSLLDRERGRSPDRRQDSRYRSDYDGRTPSRHRSYERSERERERHRRHNDNR 299
Db 546 RKGSSPS-SRHQAKARSPVRRSP-----TPVNRRS-RRSSASARSDDRR--RR 590
QY 300 SPFLERSYKKEKKRSGRSTGLSVPEPAGCTPDLPGELITKNTDSMAPPLEIVNHRSPS-R 358
Db 591 SPSSSSPSR-----SRSP--VLHRSPPSR 614
QY 359 EKKRARWEEKDRMSDNO---SSGDKNTSTIKEKEPEETMDK-----NEEEEEE 406
Db 615 GRKHORERRSPGRSLSEODRVONSKLLKRTSVPTDKRKQLEKLLGVGRVETKQDERK 674
QY 407 LKRPWIRCTHSENYSSDPMDOVGSDTVVGTSLRLDYDKFEELGS-----454
Db 675 SDKLSKRSVYRHHGSGMSVENSSEGRSPVSSKVKD-SEQVEKENSDLDANLSCDSKD 733
QY 455 ---ROEKAKAARPPWEPKTKLDEDESSSECESEDSTCCSSSDSEVFDVIAETIKRK 511
Db 734 TIRHQIKDKNRR-----KNK-----RSSREVSDD-----NGSSDSV-DDRKEAKRR 776

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 18:20:28 ; Search time 54 Seconds
(without alignments)

2444.940 Million cell updates/sec

Title: US-09-900-425a-2

Perfect score: 7500

Sequence: 1 MMGNTCHRMSPHPRGCP...MKMEREHQREPPEDTEIDK 1374

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: PIR-71:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	890	11.9	412	2	T21419 hypothetical prote
2	558	7.4	682	2	T21420 hypothetical prote
3	316	4.2	551	2	SS7447 HpbR1-7 protein -
4	315	4.2	249	2	B69693 ribonuclease III (
5	307.5	4.1	229	2	AE1300 ribonuclease III h
6	300	4.0	229	2	AE1672 ribonuclease III h
7	293	3.9	1560	2	T42727 proliferation pote
8	285.5	3.8	225	2	B82073 ribonuclease III v
9	282	3.8	263	2	A83961 ribonuclease III r
10	281.5	3.8	232	2	F98012 ribonuclease III (
11	279.5	3.7	232	2	H95144 ribonuclease III (
12	276	3.7	948	2	A57640 retinoblastoma bin
13	275	3.7	760	2	T06291 extensin homolog T
14	274.5	3.7	1531	2	T48946 hypothetical prote
15	273.5	3.6	1006	2	G86292 hypothetical prote
16	273	3.6	891	2	G84693 probable proline-r
17	271.5	3.6	620	2	S06733 hydroxyproline-ric
18	270	3.6	226	2	B84959 ribonuclease III (
19	268.5	3.6	231	2	B86725 ribonuclease III (
20	267	3.6	224	2	A81260 ribonuclease III (
21	267	3.6	243	2	D89896 RNase III (importe
22	267	3.6	998	2	T30930 hypothetical prote
23	265.5	3.5	1002	2	T43236 carboxypeptidase C
24	262	3.5	226	1	NRECC3 ribonuclease III (
25	262	3.5	226	2	A91058 RNase III (importe
26	262	3.5	226	2	F85902 RNase III, ds RNA
27	259	3.5	322	2	S25299 extensin precursor
28	258	3.4	231	2	S76204 hypothetical prote
29	257.5	3.4	272	2	T35656 ribonuclease III -

30	256	3.4	246	2	H70187 ribonuclease III (
31	255.5	3.4	1638	2	A42091 transcription acti
32	255	3.4	226	2	AC0829 ribonuclease III (
33	254	3.4	1611	2	T38236 hypothetical prote
34	253.5	3.4	368	2	C29356 hydroxyproline-ric
35	253.5	3.4	1048	2	T31425 C-terminal domain
36	252	3.4	226	2	AF0331 ribonuclease III (
37	251.5	3.4	240	2	F64602 ribonuclease III -
38	251.5	3.4	1201	2	G84441 unknown protein (l
39	250.5	3.3	727	2	C84534 hypothetical prote
40	249	3.3	225	2	T09964 extensin Cys15 pre
41	249	3.3	907	2	E96636 hypothetical prote
42	249	3.3	1877	2	T21861 hypothetical prote
43	245	3.3	230	2	G97115 dsRNA-specific rib
44	244.5	3.3	489	2	T11622 extensin class 1 p
45	243.5	3.2	708	2	D96711 hypothetical prote

ALIGNMENTS

RESULT 1

T21419

hypothetical protein F26E4.13 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T21419

R:Lightning, J.

submitted to the EMBL Data Library, October 1996

A:Reference number: 219419

A:Accession: T21419

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-412 <Full>

A:Cross-References: EMBL:Z81070; PIDN:CA803005.1; GSPDB:GN00019; CESP:F26E4.13

A:Experimental source: clone F26E4

C:Genetics:

A:Gene: CESP:F26E4.13

A:Map position: 1

A:Introns: 359/2; 400/3

Query Match 11.9%; Score 890; DB 2; Length 412;
Best Local Similarity 43.1%; Pred. No. 3e-40;

Matches 181; Conservativity 85; Mismatches 136; Indels 18; Gaps 7;

QY	944	TLINISRLGDDPTPSRNHNERLEFGDAVPELTGSHLYLRLPSLEGGATYRTAI	1003
DB	2	SLFNIRKGTSGGP---ILHNERLEYGDVAVELIVSHILYEMLTGHEGGATYRTAL	57
QY	1004	VONQHLAMLAKKLEIDPFMLVAHGPDLCRESDLRHAMNCFEALIGAVYLEGSLBEAKOL	1063
DB	58	VQNRNATLAKNCRIDEMQYSHGADLVNAEFKALNANFAVMAALYIDGGLAPCVI	117
QY	1064	FGRLTF-NDPDLREWNLPLRPLDLOEPNDRQLETSPLYQKLTREBEAIGYFTYVR	1122
DB	118	FSKAMGHQPVLEKEKWDHINEHEIKREDPOGDQDSFITPILSTFHALEERIGIOFNIR	177
QY	1123	LARATLTVGPNHTLICHNORMERLGSIMQIVATEVLFHFPDHHGHGHTLRSSLY	1182
DB	178	LAKATATRNIPINDTLTKHNRLEWLDGVSVDLVSDLYRRFPYHNEGHMSLRTISLV	237
QY	1183	NNRTOAKVAEELGMOEYATVNTKTRPV---GLRTKTLADLLESFIALLYTDKDLVYHT	1239
DB	238	SNQTOAVVCDDLGTFEFLV-----KAPYTPPELKLKDKADYEAIRIGALYVDRGTEHCRA	292
QY	1240	FMNVCFEPPLKEFLINQDNNDRKSQLOQCCLTLR-TEGNEPDIPLYKTLQTVGSPHARTY	1298
DB	293	FIRIVCPRLKRIIESEKNNDAKSHLOQWCLAMPSSSEPPMPERYVLGEGPNNRIF	352
QY	1299	TVAVYFKGRIGCGKSPSIQQAEMGA---AMDALEKYNPPQW-AHQKRFIGKRYQOEKE	1354
DB	353	KIAVYKGRKLASAEISNVHKAELRYAELALANLESMSFSKMAKNNNSMRRRLDQTS	412

RESULT 2
T21420
hypothetical protein F26E4.10 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T21420
R:Lightning, J.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19419
A:Accession: T21420
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-682 <WIL>
A:Cross-references: EMBL:Z81070; PIDN:CAB03006.1; GSPDB:GN00019; CESP:F26E4.10
;Experimental source: clone F26E4
C:Genetics:
A:Gene: CESP:F26E4.10
A:Map position: 1
A:Introns: 119/3, 226/2, 303/2, 377/3
C:Superfamily: *Caenorhabditis elegans* hypothetical protein F26E4.10

	Query Match	7.4%;	Score 558;	DB 2;	Length 682;
	Best Local Similarity	24.1%;	Pred. No. 2.7e-22;		
	Matches 178;	Conservative 126;	Mismatches 296;	Indels 140;	Gaps 23;
QY	277	RHSYSERREPRERHHRHNDNRSPDLESYKKEKRSGRSYGLSVPEPAGCPRLP-G	335		
Db	14	KHKRRKKKKQKQYERHKEEMOOLGRFQNPSTSSAPPDYKIKPLPRESTALPFG	73		
QY	336	ELIKNTDS-----WAPLEIVNHRSPSRKKRARRAMEEKDRMSDSSGKDKNYSIK	368		
Db	74	DSRLTEKDYETNYMDLP--VYSTHSALIK-----SNRVYIK	110		
QY	389	EKEPEETMPDKNEEBEELK-----PWIRCTHS-----	418		
Db	111	AEEKEKMKIMIKASTTSLDLPQFKILLETYKTRRLADADVPYIIHCHSCGKKRTPKOK	170		
QY	419	---ENYSSDPMPQVDSITVVGTSRLDLYDFEELGSRQEKAKARPP--MEPPKTL	473		
Db	171	GGSESTASVSDSDMS-----QDEASTEPFNROAPRDKT	208		
QY	474	DEPLESSSESECDSDSTCS-----SSSDSEVFVDIAEIKRKAHPDRH	519		
Db	209	GE-----VKDEKOTCNRRNOQRAKRLRNPKEERQITLLKGGIDRKKTHPGH	258		
QY	520	DELYNDPQGMNDGRLCSAKARRTGIRHSYIPEEAIKRCRMTNNAGRLHY--RIT	577		
Db	259	PDLSFNEKGGIAGNRCPCRCRPERIKCGIKHGVIYEDKADICK--KNGENLHYUULLRYT	316		
QY	578	VSEPTFLNDPRVILEYDDHEYIFEGFGFMFANAPL---TNIPLCYKTRINITYTHTE	633		
Db	317	PLPSEMOI--YRHNMAINGEEFEFGFSLITHAPLRDGMTAPRLCK---YSMDIEQLVE	371		
QY	634	EMMP--ENFCVKGLEFLSFLFDLILELYDMNLKGRLEDSPPCCRPREHMFREVRPLDG	692		
Db	372	EFPMDCECFDEDCDMLFEYIFHEIRFEMDLELRPHKRLPSDVESCPIHIMPRVQ--TKDD	430		
QY	693	GKVLVMOHILLYL--RCSKALVPEEELIANNLOMBELEMOKYAEBCGMVYTPRGKPPSS	751		
Db	431	LYOIMSSKTVLAFETSSKSGSEITMSDEPNRLCDAODQFTRNTSKHKQSLVLTNKKRPA	490		
QY	752	VRIDQDRQFNDVITPPIIVHFGIRPAQSLVAGDPDPOYOKIMKSYKRLNHLLANSPKY	811		
Db	491	IRADWEERDEBEKEV---YVHNMAIRQYLTATSLSPRIAFLEKTLINKMIO--EKQSSGY	545		
QY	812	QIDKQKLAOREALIKROKNTMRRETVLSSQFMKTKGTSIDVCONAHMLVLTNHR	871		
Db	546	NKDFEK---TKNELEHIKRENSAENILKLRPVAGFLETGKLRVAAHVMTILACHNR	602		
QY	872	YHDCSLMHKILGYFQORCLDQALMTNPSHNLNFGANPNRHAKRSLNSGCIROPKYGDK	931		
Db	603	YNSLSDVEEVLLEYFKNDRVYETLALMSSKSKSYGPRIDHVKMIMTINCYGR--KYG---	658		

```
QY 932 VHHMHRKKGINTLINIMSR 951
      :|:| : : :
Db 659 AEDKREKKRGSRLKIKFLDK 678
```

RESULT 3
S57447
HPBRI7-7 protein - human
N:Alternate names: HPBRI7-4 protein
C:Species: Homo sapiens (man)
C:Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 24-Sep-1995
C:Accession: S57447, S57489
R:Flieschauer, K.L.
submitted to the EMBL Data Library, June 1992
A:Reference number: S57447
A:Accession: S57447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-551 <FL>
A:Cross-references: EMBL:X67336; NID:g871300; PIDN:CAAA7751.1; PID:g871301
A:Accession: S57489
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-551 <FL2>
A:Cross-references: EMBL:X67337; NID:g871298; PIDN:CAAA7752.1; PID:g871299
C:Genetics:
A:Introns: 231/3
C:Superfamily: unassigned ribonucleoprotein repeat-containing proteins; ribonucleopro
;82-151/Domain: ribonucleoprotein repeat homology <RMA4>

```

Query Match          4.2% Score 316; Db 2; Length 551;
Best Local Similarity 28.0%; Pred. No. 1.6e-09;
Matches 111; Conservative 32; Mismatches 106; Indels 148; Gaps 21;

QY 19 PRGRGAGCARPSAASFRRQNLRLHPQQPVQYGERPSAPSTTFSSNPAFNF---LPP 74
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 198 PGGGKRGKRFGGA-----VPGGDHFRPAGP-----GGPRPPFPAGGTTP 237

QY 75 RRDVVFFRRPMPRPSAQGLRPCTLRPP----FRNNQMRNHPPV-----PPC--FPP 119
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 238 RP---PLGPBPBRPPRPPGGQVLPRPLAGRNRRGDRPPRVLPRLPQQPFGQRLPGLRP 294

QY 120 MPPEMPCPNNNPVVGARPGQ-----GTFP-----FMMPSPSMRPP--PPV 159
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 295 GPPE-PVGVGYRPPRPPRPPQGQRRPPRRPFRRRRRPPGLRLTLARPHLLGRRPGARRP 353

QY 160 MPQOVNYQYPFGYGHNNFPSPFSFNQNNBSGLTLPANSSSP--HRRLPPV----- 210
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 354 AP-FVN-----PAF-----FPPPT-----NSGMPTSDSRGPPPTDYGRRPPPDGDDYG 396

QY 211 PLKPASRRRSRPERLKTHDDRHRRHS----- 237
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 397 PRGEMDTARKLPLEAEFEELIMNRKNRAISSAISRAVSADASAGYSAILETVTAISLIK 456

QY 457 QSKVASDDRCKLVLISSLDQCLHGIESKSYGGSGSRRERSRRERDSHR--SRKSRRHKRSR 514
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 515 DRHDYYREKRSRERERRRRDRD--RDORERDREREYR 549

RESULT 4
B69693
ribonuclease III (EC 3.1.26.3) - Bacillus subtilis
N:Alternate names: RNase O
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C:Accession: B69693; J04821
R:Kunstl., F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
Z.; Bron, S.; Brouillett, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;

```


A.: Ehrlich, S.D.; Emmerson, P.T.; Entlian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier,
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koehler, P.; Konningstein, G.; Kirogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, Y.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schreuter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot
akech, M.; Tamakoshi, A.; Tanaka, T.; Teppstra, P.; Tognoli, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A63580, MUID:980404033
A:Accession: B69693
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-249 <KUN>
A:Cross-references: GB:299112; GB:AL009126; NID:g2633902; PIDN:CAB13466.1; PID:g2633965
A:Experimental source: strain 168
R:Oguro, A.; Kakeshita, H.; Takamatsu, H.; Nakamura, K.; Yamane, K.
Gene 172, 17-24, 1996
A:Title: The effect of Srb, a homologue of the mammalian SRP receptor alpha-subunit, on
A:Reference number: J04819, MUID:96257247
A:Accession: J04821
A:Molecule type: DNA
A:Residues: 1-78, 'P', 80-137, 'K', 139-170, 'P', 172-242, 'E', 244-249, 'LNPPYDSGPFQYVCLRI' <GCU
A:Cross-references: DDBJ:D04116; NID:g1389548
A:Note: The sequence in Genbank entry D04116, release 106.0, (PID:g1389549) has 138-Glu
A:Gene: rncs; srb
A:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology
C:Keywords: hydrolase
F:172-245/Domain: double-stranded RNA-binding repeat homology <DSR>

Query Match 4.2%; Score 315; DB 2; Length 249;

Best Local Similarity 34.7%; Pred. No. 6.8e-10;

Matches 85; Conservative 46; Mismatches 96; Indels 18; Gaps 7;

QY 1104 LQKTEFEAIGVFTYVRLARFRTLRVGFNHL--TGH--NORMEFLGDSIMQVATEY 1161
DB 17 VEOKEERSEIVHFOEKELLYQAFTHSSVYNEHRRKRYDENLEFLGDAVLELTISR 76
QY 1162 LFIHPDHGHTLLRSSLVNNRTQAKVAEELGMOEYATINDTKRPVGL--TLADL 1220
DB 77 LFAKYPAMSGDLTKRAALYCEPSLVSLHLSFGDLVLLGGEEMTGGRKRPALLADY 136
QY 1221 LESFIALYDKDLEVYHFMNCFEPRLKEFLINQDMPKSQLQCCLTLETEGK-- 1277
DB 137 FEAFIGALYLDQGLEPESFLKYVFPKINDGAFSHVM--DFKSQLD--YVORDGKSL 192
QY 1278 EPDIPKLTQTVGSHARTYVAVYFKGERIGCGKPSIQQAEMGAAMALEKYNPQM 1337
DB 193 E----KISNEKPAHNRFEALVSLKGEPLGVNGRSGKAEQAHQAQALAKL--QK 244
QY 1338 AHOKR 1342
DB 245 HHTKQ 249

RESULT 5
AEI300
ribonuclease III homolog rncs [imported] - *Listeria monocytogenes* (strain EGD-e)
C:Species: *Listeria monocytogenes*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
C:Accession: AEI300
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefl, J.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maltournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A:Title: Comparative genomics of *Listeria species*.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AEI300

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-229 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CAC99883.1; PID:g16411259; GSPDB:GN00177

A:Experimental source: strain EGD-e

A:Gene: rncs

C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

Query Match 4.1%; Score 307.5; DB 2; Length 229;

Best Local Similarity 33.0%; Pred. No. 1.5e-09;

Matches 76; Conservative 51; Mismatches 86; Indels 17; Gaps 7;

QY 1109 EPEEAIGVFTYVRLARFRTLRVGFNHL--TLGHQRMFLGDSIMQVATEYLFHF 1166
DB 6 ELQESVGFDFKVVLELQQAFTHTSSVYNEHRRNVKRDNELEFLGDAVLELVSDYLFKY 65
QY 1167 PDHHEGHTLLRSSLVNNRTQAKVAEELGMOEYATINDTKRPVGLRTR--TLADL 1224
DB 66 PDMAEEMTKMRAALYCEPSLVFEPAVHFSTY--VRLGGEKAGGRTRPALADVFESF 124
QY 1225 IALYTDKDELYVHFEMVNFEPRLK--EFLINQDMPKSQLQCCLTLETEGKEDIP 1282
DB 125 ICALYDNGIDKRVYFLERVFPKIDAGAYLQTVDY---KTQLQELIV-----QDRDYL 175
QY 1283 L-YKLTQTVGSHARTYVAVYFKGERIGCGKPSIQQAEMGAAMDALEK 1331
DB 176 IEXDILGETGPAHNRFAQVAVYNGVGLGSGRTRKQAEQSAQAQFAINK 225

RESULT 6

AEI672
ribonuclease III homolog rncs [imported] - *Listeria innocua* (strain Clp11262)

C:Species: *Listeria innocua*

C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C:Accession: AEI672

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entlian, K.D.; Fsihl,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001

A:Authors: Krefl, J.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maltournam, A.;
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla

A:Title: Comparative genomics of *Listeria species*.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AEI672

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-229 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97149.1; PID:g16414420; GSPDB:GN00178

A:Experimental source: strain Clp11262

A:Gene: rncs

C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

Query Match 4.0%; Score 300; DB 2; Length 229;

Best Local Similarity 31.8%; Pred. No. 3.8e-09;

Matches 76; Conservative 53; Mismatches 88; Indels 22; Gaps 8;

QY 1109 EPEEAIGVFTYVRLARFRTLRVGFNHL--TLGHQRMFLGDSIMQVATEYLFHF 1166
DB 6 ELQESVGFDFKVVLELQQAFTHTSSVYNEHRRNVKRDNELEFLGDAVLELVSDYLFKY 65
QY 1167 PDHHEGHTLLRSSLVNNRTQAKVAEELGMOEYATINDTKRPVGLRTR--TLADL 1224
DB 66 PDMAEEMTKMRAALYCEPSLVFEPAVHFSTY--VRLGGEKAGGRTRPALADVFESF 124
QY 1225 IALYTDKDELYVHFEMVNFEPRLK--EFLINQDMPKSQLQCCLTLETEGKEDIP 1282
DB 125 ICALYDNGIDKRVYFLERVFPKIDAGAYLQTVDY---KTQLQELIV-----QDRDYL 175
QY 1283 L-YKLTQTVGSHARTYVAVYFKGERIGCGKPSIQQAEMGAAMDALEKYNPQMAHQ 1340

Db 176 IEYDILGETGPPAHNNKAFDAQVIVNGVIGKGGRTKKAQEOSAOFATIN-----OLTHR 229

RESULT 7

TA2727
Proliferation potential-related protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000

C:Accession: TA2727

R:Witte, M.M.; Scott, R.E.

Submitted to the EMBL Data Library, November 1998

A:Reference number: Z22246

A:Accession: TA2727

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1560 <WIT>

A:Cross-references: EMBL:U83913; NID:g3858884; PID:g3858885; PIDN:AAC72432.1

A:Experimental source: strain Balb/C

C:Genetics:

A:Gene: P2P-R

C:Function:

A:Description: involved in hnRNP association and Rb1 binding

C:Superfamily: RING finger homology

F:57-107/Domain: RING finger homology <RRN>

Query Match 3.9%; Score 293; DB 2; Length 1560;

Best Local Similarity 23.8%; Pred. No. 1e-07;

Matches 151; Conservative 72; Mismatches 200; Indels 212; Gaps 34;

QY 25 HGARPSA--PSFRPONLRLHPQDPVOYQYEPSPA-----PSTTFSNPAFNLFP 73
 Db 296 NARBGSGRPGWGHSHN-KLGYLVSPQOIRRGERSGYSINGRHNSRSQRTQGP-LP 353
 QY 74 PRDPFVPRPMPSPAQGLPCPLRPPIRPPRNQHMPPRPVPCFRMPMPMCNNPPVP 133
 Db 354 ARPVVPPVPP-----PLKPPP-----HTLPLP--PGVLPQPSQGP--P 391
 QY 134 GAPPGGCTFFPMPPSPMHP-----PPVMPQOVVY----- 166
 Db 392 GQPPAG---YSVPPGPPAPANISTACFSGVPTAHSNTPPTQAVLLSEEFYRQK 448
 QY 167 -----QYR--PGYSHHFPSPNSFQNNSSFLPSANSSSHFHNLPPYLPKAPS 217
 Db 449 RUKESKPEYSGSSYSRSSYTDKSRSGSTRSRSRSRSHSRSHSRSPYP----- 502
 QY 218 ERRSPERLKHNDHRRHDSHGRHSRLDRERGSPPDRRDSRYRSDYD--RGRTPS 276
 Db 503 -RRGKGSKNY---RSRSKSHG---YHRS-----RSRSPPIRRYRSRSRSPQAFRGOSPT 550
 QY 277 RHRSTYRSRERERER-HRHR-----DNRRSPSLERSYKKEYYK 313
 Db 551 K-RNVPRG-ETREYEFNRYREVPYDIKAYYRGSVDPRDPPEKRYREWEKRYWEK 608
 QY 314 SGRSGLSVPPPACTPELPGELIKNDSNAP-----FLEYVNHSPREKRRARWEK 369
 Db 609 YKGAVAQAQPPSA-----NNEDESPERLLPLNIRN--SPTRGRREDIYAGQ 655
 QY 370 DRMSDN-----OSSGDKMYTSTIKEREPEETMPD-----KNE--EE 403
 Db 656 SHRNNNLGNVPEKSLTSDSHNAKNPASKESLENVGGDKGNKHKHRRRRDEKCEE 715
 QY 404 EBEELKPMWI---RCHSENY-----YSDPMDQVGDSTVVGTSR 440
 Db 716 SESFNPPELLGFRKRCSSGIDETKTDTLFFPSREDATPVRODPMO--AESITFKVS 773
 QY 441 LRDLYDKREBELGROEKAKA-----ARPPWPKTKLDEDLSSSEF---CES 487
 Db 774 DKDKREKPKVKSDKTRKSGSATAKDNVLFKSKGPOEKVDGDRKSPSEPLKKA 833
 QY 488 DEDST-----CSSSDSEVFVDIAEIKRRKAH 514
 Db 834 KEATKIDSVKRSSSSSQDEKVTG-----PRKAH 863

RESULT 8

B82073

Ribonuclease III VC2461 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: B82073

R:Heidelberg, J.F.; Eison, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.

chardson, D.; Ermolaeva, M.D.; Yamatyan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of Both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: B82073

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-225 <HEI>

A:Cross-references: GB:AE004316; GB:AE003852; NID:g9657034; PIDN:AAF95603.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2461

A:Map position: 1

C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

Query Match 3.8%; Score 285.5; DB 2; Length 225;

Best Local Similarity 32.2%; Pred. No. 2.2e-08;

Matches 75; Conservative 42; Mismatches 101; Indels 15; Gaps 5;

QY 1101 SPVLQKLEFEAEAGVIFTHVRLARAFTLRVGFNHLTGNHQMFEGLDSIMQVATE 1160

Db 2 TPVNMKLT---SKRGYFKETELNLTATFRSANGK-----HNERLEFGDSTLSFVIAD 53

QY 1161 YLFHPEPDHGHGHTLIRSSLVNNRQAKYAEELGQGEVAFITDKRKRPVGTGK-LAD 1219

Db 54 ELVRRPKNVGGDSRRKATLVKGNLTAEIGREFDLDGYTKLGPGLKSGGFRSDILAD 113

QY 1220 LLESFIALYTDKDLVHVFHMVNCFFPRKPEFLINDQNMDPKSQLQOCCLTLRTEGKEP 1279

Db 114 AVALIGATVLDSPLETFARSTVLEWYHGRLEIKPGASQDPRTRLOEF-----LGRKK 168

QY 1280 DIPYKTLQVGFSHARTTVVANYFKG-ERIGCGKPSIQOAMGMAADALEK 1331

Db 169 PLPVYTVTNIGAHNOEFTVACEVAGMDPVVIGKTSRRKAKQAATAETALDO 221

RESULT 9

A83961
Ribonuclease III rncs [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C:Accession: A83961

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: A83961

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-263 <STO>

A:Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BA06208.1; GSPDB:G

A:Experimental source: strain C-125

C:Genetics:

A:Gene: rncs

C:Superfamily: ribonuclease III; double-stranded RNA-binding repeat homology

Query Match 3.8%; Score 282; DB 2; Length 263;

Best Local Similarity 30.9%; Pred. No. 4.1e-08;

Matches 77; Conservative 50; Mismatches 110; Indels 12; Gaps 5;

QY 1084 HPIQLOEPNDRQILIESPVLOKLTGEFEAEAGVIFHVRLLARAFTLRVGFNHN--LTIG 1141

Db 15 HSERROP---KRLTLFAKQOQMFDELRLTLNLFANKRLVQAFHSSVYVNEHRIQSK 71

Db 472 LKEEKKKKKLDELFTNDAFKELMEYKKIÖKERRRSFSRSKSPYSSGSSYSRSSSYTSKSGS 531

QY 185 FQNNPSSFLPSANNSSSPFPHLLPPYPLPKAPSERSPERLKHDDHRRHSHGGERH 244

Db 532 GSTRSHSYRSRFSRSHSRYSRSPFP-----RGRKSKSNY---KSRSHG---YH 578

QY 245 RSLDRERGRSPDRRRÖDSRYSDVD-RGRTPSRHRSYERSRER-----289

Db 579 RS-----RSRSPPYRRYHRSRSPQAFRCÖSPNKRNVPGÖGETEREYFNRYREVPYDMK 633

QY 290 -----ERHRHÖNRRSPLESRYKKEKRSRSYGLSVPEPPACTBELGE 336

Db 634 AAYGSRVDRDPPEKERYE-----WEKRYREWEKYYKGVAAGQPPSA-----679

QY 337 IIKNTDSMAP-----PLEIYNHRSPSREKRAWEKEKDRSWNÖSSG-----KD 381

Db 680 ---NRNFPSPERPLPLNTIN--SPTRGRREDYVGÖSHRSRNSNYSNPEKLSADGHNÖ 734

QY 382 KNYTSIKEKEPEETMPD-----KNEEBEELKPYWIRCTHSENYSSDPM--427

Db 735 KÖMTKSEKERESEVAPGÖGKNKKHKKRRKKEESESGFLNPILLETSRK---SREPTGV 790

QY 428 -DQVGSYTVGTSR-----LRDLYÖKEEE---LGSÖKAKARAPWMPPTKIDEDLE 478

Db 791 EEKKTÖSLFVPSRDDATPVRO--EPMDAESITFKSVSEKDRER---DKPAKGDÖTKR 845

QY 479 SSSESECEDEDSTCSSSDSE-----VFDVIAEIKRKAHPDRL 518

Db 846 KNDGSAVSKKENIVKPAKPGÖKGVGDVRLDLMNÖLKKPKRRRL 893

RESULT 13

T06291
exensin homolog T9E8.80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-Jun-1999
C:Accession: T06291
R:Bevan, M.; Pohl, T.; Weizenegger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schuel
submitted to the Protein Sequence Database, April 1999
A:Reference number: 215588
A:Accession: T06291
A:Molecule type: DNA
A:Residues: 1-760 <BEV>
A:Cross-references: EMBL:AL049608
A:Experimental source: cultivar Columbia; BAC clone T9E8
C:Genetics:
A:Map position: 4
A:Note: T9E8.80

Query Match 3.7%; Score 275; DB 2; Length 760;
Best Local Similarity 35.1%; Pred. No. 3.8e-07;
Matches 88; Conservative 15; Mismatches 96; Indels 52; Gaps 12;

QY 14 PGNCGRGNGGARGASAFSRQNL-----RLHPOÖPYÖYÖEPSSAS-----60

Db 361 PGNPAÖRSÖGÖKALFSR---PVNCGSFSGSRVSP--RPVVTPLRPPLSPSPPPAP 416

QY 61 ----TTFNSPAPNLFPPRPF-VPPPPMPPSAÖGLPÖPÖIRPPFPHÖMRHFPYVP 115

Db 417 FSTPPLTLPPPPS---PPPPVYSPPPPPPPPPVYSPPPPPPPPPPPPPPPPPPPPP 474

QY 116 CEPFM---PPMPCPNRPVPGARPGÖGTFRPMRPPSMRHPPPV-----MPO 162

Db 475 PPPPVYSPPPSPPPPPVYSPPP---PPPPPPVYSPPPPPVYSSPPPPSPAP 530

QY 163 QVNYÖYÖPÖGYSHNPPPPPFNSFÖNNP--SFLPSANSSSHF---RLHPLYL-----212

Db 531 PVTCTRPAPPPHSPPPQFSPPPEPYYSPPRPHSSPPRPHSPPPRPPYPLYS 590

QY 213 -PKAPSERSP 222

Db 591 PPPPPTPVSSP 601

RESULT 14

T48946
hypoetical protein T15B3.60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T48946
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S
submitted to the Protein Sequence Database, April 2000
A:Reference number: 225009
A:Accession: T48946
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1531 <ÖÖR>
A:Cross-references: EMBL:AL163975; GSPDB:GNO0061; ATSP:T15B3.60
A:Experimental source: cultivar Columbia; BAC clone T15B3
C:Genetics:
A:Gene: ATSP:T15B3.60
A:Map position: 3
A:Introns: 45/2; 100/3; 183/3; 227/3; 278/3; 322/1; 421/3; 475/2; 584/1

Query Match 3.7%; Score 274.5; DB 2; Length 1531;
Best Local Similarity 21.3%; Pred. No. 9.7e-07;
Matches 247; Conservative 142; Mismatches 420; Indels 351; Gaps 58;

QY 372 WSDNÖSSGÖKKNYTSIKEPEETMPDKNEEBEELKPYWIRCTHSENYSSDPMÖYÖ 431

Db 455 YSÖSQKHAKÖSNKSIMFL--ERGNPKÖRDHLDMRREVLÖDPEAPNLSÖPPYKNG 512

QY 432 DST-VGTSRLRDLÖYKFEELSGRÖKAKARPWE---PP--KTÖDED-LESSSES 483

Db 513 HGVEIKESMWTIPDSNITVSEBASTÖTMSD--PPSRNEÖLPÖCKKRLDNNLÖSNGRE 569

QY 484 ECEDEDSTCSSSDSEVPVIAEIKRKAHPDRLDEL---VYNDPGÖMNDP-----534

Db 570 KVASSKSSSSSAGSK-----KKELHGTTCANALSÖT---GENIDATÖYÖK 617

QY 535 ---LÖKCSAK-----ARRTG-IRHSYÖGEÖK---PCRP-----564

Db 618 FDFCINISGEYSSFLSLSTLAEÖGVEMÖDLYVRLKVASVSPCÖGIRLSÖELV 677

QY 565 -----TNNAGRLPHRYTYSPPNFLTDRPYEYÖDHÖYIEÖGFSMFPAHLTNI 615

Db 678 AKYÖÖFFENGMÖFKLFGVSKSÖGTÖKREFLÖ-----TÖTSSLMPAPMFLLPVETN 730

QY 616 PLCKVIRPNIDYTHFTEEMPEVFCYKÖLEFSLFÖRD-IÖELYDMNKGFLPÖDSP 674

Db 731 DLAS--SATIDMS-----AINSCASIVE---FLKNSLÖDLRD-----SDGNG 768

QY 675 CÖPRFHMPRÖVRLPDGÖKEVLSMÖLÖLLYLRCSKALVPEEELANMÖEELMÖKYA 734

Db 769 C-----NTSSGÖEVL-----LÖDKMEETNLIHFANAÖSDKNS 800

QY 735 EÖCKGMVTPNGTÖPSSVRI--DÖLDRÖQNPÖYÖTPIIVHÖGIRPÖLÖSVAGDÖYÖK 792

Db 801 LÖELVLAIHÖGRYLSIVEAVSÖSANSÖPREVD-----ASSGÖYVÖAYÖEN 846

QY 793 LÖKSY-VKLRLHLANSÖ--VYKÖTÖKÖ---LÖRÖEALÖKIRÖK--NTÖRRRÖVÖELSS 844

Db 847 --KRYGIVLÖH--PÖNÖLMLKÖSHNHNLLÖDENEMVYKTERKAGNÖVKKRÖPNI----898

QY 845 ÖGFÖKRTGIRSDVÖCHAMMLÖVLTHNIRYÖNÖCLNHLÖLIGYÖTÖDRCÖLÖLÖMTÖPNS--H 902

Db 899 -----HAILPPEL-----LÖRID-----VPRVLYKSIYLLPSVNH 928

QY 903 HLNÖGMPDÖHÖRNSLSCÖRÖYÖGÖRKVHNHMKKÖGINTLÖNMSRLÖGÖDDPÖPSRK 962

Db 929 RLÖSIMLÖSÖLÖREI--DÖSÖDNFSISÖTÖ---LEÖVTTL-----CÖPS 969

QY 963 NHNÖRLEFÖGÖVÖEFLSVLYLÖPÖSLÖEGGÖLATYÖALYÖNÖHLÖML--AKÖLE---1017

Db 970 FSWÖRLELÖGÖSVLÖKYVÖASHLÖPLKYÖDKÖGÖLSRÖRÖSÖTISNSNHLRLLTÖSRKÖGÖY 1029

```

QY      161 PQQVNTQYPPGYSHHNPPPPSFGNSQNNPSSFLPSAANNSSSPHFRHLP---YLPKAPS 217
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      197 GDDTQ-QPPPLW---LPPPPFGD-ETPPPVFSLPPLDEPP---MPTTWLPDDVA 246
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      218 ERSRPERLKH-----YDDHRRHDSHGSGEHRSLDRREGSRSPDRRODS 265
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      247 QTSSAAEQDLPPLVTITTEALENPNSHRHD-----ENKGLDKRRNP-RVKSRRRSR 299
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      264 R 264
          |
Db      300 R 300

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2002, 17:49:53 ; Search time 88 Seconds
(without alignments)
2701.080 Million cell updates/sec

Title: US-09-900-425a-2
Perfect score: 7500
Sequence: 1 MMQNTCHRMSEFHPGRCGCPR.....MRWERHOREPDEDIDIK 1374

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp Unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7486	99.8	1374	4	O9NR4
2	4071	54.3	769	4	O9NW73
3	2963	39.5	1327	5	O9XYN5
4	2847	38.0	541	4	O9Y2V9
5	2811.5	37.5	1071	5	O960Y4
6	1973	26.3	380	4	O9Y4Y0
7	1448	19.3	1086	5	O01327
8	1417	18.9	860	5	O9U908
9	711	9.5	134	11	O9CTG2
10	340	4.5	1151	5	O9YAV4
11	316	4.2	551	4	O16630
12	306.5	4.1	588	4	O9BW18
13	300	4.0	229	16	O9ZAK3
14	288.5	3.8	1591	11	P97868
15	285	3.8	409	10	O9SAM1
16	285	3.8	478	4	O9BSJ7

17	285	3.8	652	5	O95TS9	O95TS9 drosophila
18	284.5	3.8	1616	4	O96PH3	O96PH3 homo sapien
19	282	3.8	263	16	O9KA05	O9KA05 bacillus ha
20	279.5	3.7	232	16	O970G6	O970G6 streptococ
21	276	3.7	948	4	O152S0	O152S0 homo sapien
22	275.5	3.7	594	5	O9VEP4	O9VEP4 drosophila
23	275	3.7	760	10	O9TOK5	O9TOK5 arabidopsis
24	274.5	3.7	1531	10	O9LXW7	O9LXW7 arabidopsis
25	273.5	3.6	1006	10	O9LMO1	O9LMO1 arabidopsis
26	273	3.6	891	10	O9ZM08	O9ZM08 arabidopsis
27	273	3.6	894	10	O9FTB2	O9FTB2 arabidopsis
28	272.5	3.6	702	5	O9VSH4	O9VSH4 drosophila
29	271.5	3.6	230	16	O9A105	O9A105 streptococ
30	271	3.6	2994	5	O95ZG5	O95ZG5 dictyostell
31	269	3.6	884	4	O00302	O00302 homo sapien
32	268	3.6	243	16	O931T1	O931T1 staphylococ
33	267	3.6	224	16	O9PM40	O9PM40 campylobact
34	267	3.6	243	16	O990N7	O990N7 staphylococ
35	267	3.6	988	5	O01864	O01864 caenorhabdi
36	266	3.5	687	10	O948Y7	O948Y7 volvox cart
37	265.5	3.5	1002	3	O13849	O13849 schizosach
38	265	3.5	740	3	O96U76	O96U76 neurospora
39	261	3.5	763	2	O9XDH2	O9XDH2 mycobacteri
40	261	3.5	1236	4	O9C012	O9C012 homo sapien
41	259.5	3.5	555	10	O9FP06	O9FP06 chlamydomon
42	259	3.5	322	10	O09084	O09084 lycopersico
43	259	3.5	1217	4	O9ULI5	O9ULI5 homo sapien
44	258	3.4	1400	11	O9ESU6	O9ESU6 mus musculu
45	256	3.4	956	10	O9LJ64	O9LJ64 arabidopsis

ALIGNMENTS

RESULT 1

ID O9NR4 PRELIMINARY; PRT; 1374 AA.
AC O9NR4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE RIBONUCLEASE III.
GN RN3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20538440; PubMed=10948199;
RA Wu H., Xu H., Miraglia L.J., Crooke S.T.;
RT "Human RNase III is a 160-kDa protein involved in preribosomal RNA
RT processing.";
RL J. Biol. Chem. 275:36957-36965(2000).
DR EMBL: AF189011; AAF80558.1; -;
DR InterPro: IPR001159; DS_RBD.
DR InterPro: IPR002865; P_Rich_extensn.
DR InterPro: IPR000999; RNase_3.
DR Pfam: PF00035; dsrm; 1.
DR Pfam: PF00636; Ribonuclease_3; 2.
DR PRINTS: PR01217; PRICHEXTENS.
DR SMART: SM00358; DSRW; 1.
DR SMART: SM00355; RIBOC; 2.
DR PROSITE: PS50137; DS_RBD; 1.
DR PROSITE: PS50157; RNASE_3_1; 2.
DR PROSITE: PS50142; RNASE_3_2; 2.
SQ SEQUENCE 1374 AA; 159244 MW; 888108523912F705 CRC64;

Query Match 99.8%; Score 7486; DB 4; Length 1374;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 MMQNTCHRMSEFHPGRCGCGARPSAPSPDPNLRLLHFGQPPVQYQYEPSPAPS 60

```

Db 1 M0AGTCTARMSPHRCGRGCRGGRGAGAPASAPSRPQNLRLHLDQOPVQYQYBPAPS 60
QY 61 TTFNSAPNPLPRPDPVPPPPMPSAOGPLPPCIRPFPNHNKHPPPVPCPPM 120
Db 61 TTFNSAPNPLPRPDPVPPPPMPSAOGPLPPCIRPFPNHNKHPPPVPCPPM 120
QY 121 PPMPCPNPPVPGAPPGQGTFFPMMPSPMHPPPVPMQOVNYQYPRGYSNHPPP 180
Db 121 PPMPCPNPPVPGAPPGQGTFFPMMPSPMHPPPVPMQOVNYQYPRGYSNHPPP 180
QY 181 SFNSQNNPSSFLPSANSSSPHFRHLPPYLPKAPSRBRDELKHYDDHNRHDSHGR 240
Db 181 SFNSQNNPSSFLPSANSSSPHFRHLPPYLPKAPSRBRDELKHYDDHNRHDSHGR 240
QY 241 GERHSLDRRGRSPDRRDSRYRSDYDRGRTPSRHRYSERERERERHRDRNRS 300
Db 241 GERHSLDRRGRSPDRRDSRYRSDYDRGRTPSRHRYSERERERERHRDRNRS 300
QY 301 PSLEHSYKKEYKRSRSTGLSVBPAGCTPELPGELIKNTDSNAPPLEIYNHRSPSREK 360
Db 301 PSLEHSYKKEYKRSRSTGLSVBPAGCTPELPGELIKNTDSNAPPLEIYNHRSPSREK 360
QY 361 KRAREEKKOWSDNOSGKCKNTSTIKERPEETMPKNEEEELLKPYWICTHSEN 420
Db 361 KRAREEKKOWSDNOSGKCKNTSTIKERPEETMPKNEEEELLKPYWICTHSEN 420
QY 421 YSSDPMQOVGDSYVGTSTRLDYLDKPEEELSGROEKAKARAPMPPPKTKLDEESS 480
Db 421 YSSDPMQOVGDSYVGTSTRLDYLDKPEEELSGROEKAKARAPMPPPKTKLDEESS 480
QY 481 SESECEDEDSTCSSSDSEVFDVIAETIKRKAHPDRHLDLWYNDPQMDGFLCKCSA 540
Db 481 SESECEDEDSTCSSSDSEVFDVIAETIKRKAHPDRHLDLWYNDPQMDGFLCKCSA 540
QY 541 SESECEDEDSTCSSSDSEVFDVIAETIKRKAHPDRHLDLWYNDPQMDGFLCKCSA 540
Db 541 SESECEDEDSTCSSSDSEVFDVIAETIKRKAHPDRHLDLWYNDPQMDGFLCKCSA 540
QY 541 KARRTGHSHTYPPGEALIKPCRPMTNNGRLNHRITMSPPTNLTPRYTIEDDHEXI 600
Db 541 KARRTGHSHTYPPGEALIKPCRPMTNNGRLNHRITMSPPTNLTPRYTIEDDHEXI 600
QY 601 FEGFSMAHAPLTINPLCKVIRFNIDYTHFIEEMMPENFCVKLELFSLEFIDILEY 660
Db 601 FEGFSMAHAPLTINPLCKVIRFNIDYTHFIEEMMPENFCVKLELFSLEFIDILEY 660
QY 661 FEGFSMAHAPLTINPLCKVIRFNIDYTHFIEEMMPENFCVKLELFSLEFIDILEY 660
Db 661 FEGFSMAHAPLTINPLCKVIRFNIDYTHFIEEMMPENFCVKLELFSLEFIDILEY 660
QY 721 NMLQWEELEWOKYAECKGMIVTNPGRKPSVSRIDQIDREQFNPDVITFPIIVHGRPA 780
Db 721 NMLQWEELEWOKYAECKGMIVTNPGRKPSVSRIDQIDREQFNPDVITFPIIVHGRPA 780
QY 781 QLSYAGDPOYOGLMKSYVYKLNHLANSPPKVQTDOKIQAQREALOKIROKNTMRRETV 840
Db 781 QLSYAGDPOYOGLMKSYVYKLNHLANSPPKVQTDOKIQAQREALOKIROKNTMRRETV 840
QY 841 ELSSGGEFKTGTIRSDYCOHAMLPVLTNHRHOCIMLMDLIGTTFDRCCLDLAMTHP 900
Db 841 ELSSGGEFKTGTIRSDYCOHAMLPVLTNHRHOCIMLMDLIGTTFDRCCLDLAMTHP 900
QY 901 SHHLNFGNMPDHARNSLNCGRKPYGDRKVHNMHMKKGINTLINIMSLGDDPPPS 960
Db 901 SHHLNFGNMPDHARNSLNCGRKPYGDRKVHNMHMKKGINTLINIMSLGDDPPPS 960
QY 961 RINHERLEFLGDAVEELTVSHLYLTPSLDEGLATYRAIYONCHIAMAKLELDP 1020
Db 961 RINHERLEFLGDAVEELTVSHLYLTPSLDEGLATYRAIYONCHIAMAKLELDP 1020
QY 1021 FMLYAHGPDLCRESDLRAMANCFEALIGAYULBESGEAKOLFGRLLFNDPDLREWLN 1080
Db 1021 FMLYAHGPDLCRESDLRAMANCFEALIGAYULBESGEAKOLFGRLLFNDPDLREWLN 1080
QY 1081 YPLHPLQLOEPNTDROLLETSPVLOKLEFEALGVITTHRLARATLTETVGFNHLT 1140
Db 1081 YPLHPLQLOEPNTDROLLETSPVLOKLEFEALGVITTHRLARATLTETVGFNHLT 1140

Db 1081 YPLHPLQLOEPNTDROLLETSPVLOKLEFEALGVITTHRLARATLTETVGFNHLT 1140
QY 1141 GHNORMEFLGDSIMOLVATEYLFHPDDHHEGLTLRLSSLVNNRTOAKVAEELGMOEYA 1200
Db 1141 GHNORMEFLGDSIMOLVATEYLFHPDDHHEGLTLRLSSLVNNRTOAKVAEELGMOEYA 1200
QY 1201 IINDKTRPVGLKTKTLADLESFIALYTDKDELEVHFMVCFPPRLKERTLNQMDND 1260
Db 1201 IINDKTRPVGLKTKTLADLESFIALYTDKDELEVHFMVCFPPRLKERTLNQMDND 1260
QY 1261 PKSOLOOCCLTIRTEGKEPDIPLYKTQTVGSHARTYVAVYFGERIGCKGSPISIOA 1320
Db 1261 PKSOLOOCCLTIRTEGKEPDIPLYKTQTVGSHARTYVAVYFGERIGCKGSPISIOA 1320
QY 1321 EMGAAMDALERYNFPQMAHOKFETGRVYMOBLEKEMWEHOREPDEDEDIKK 1374
Db 1321 EMGAAMDALERYNFPQMAHOKFETGRVYMOBLEKEMWEHOREPDEDEDIKK 1374

RESULT 2
Q9NM73 PRELIMINARY; PRT; 769 AA.
ID 09NM73:
AC 09NM73:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE HYPOHETICAL 89.7 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN 111
RP SEQUENCE FROM N.A.
RC TISSUE=EMERYO;
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Magatsuna M., Hosoi T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takasashi M., Chida Y., Ishida S., Murakami K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuno Y.,
RA Nishimura K., Iwayanagi T.;
RT NEDD human cDNA sequencing project.
RL Submitted (Feb-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AK001121; BAA91511.1;
DR InterPro: IPR001159; DS_RBD.
DR InterPro: IPR000999; RNase_3.
DR Pfam: PF00035; dsrm; 1.
DR SMART: SM00358; DSRM; 1.
DR SMART: SM00355; RIBOC; 2.
DR PROSITE: PS01317; DS_RBD; 1.
DR PROSITE: PS00517; RNase_3_1; 2.
DR PROSITE: PS0142; RNase_3_2; 2.
SQ SEQUENCE 769 AA; 89706 MW; C92261C723C047FA CRC64;

Query Match 54.3%; Score 4071; DB 4; Length 769;
Best Local Similarity 99.5%; Pred. No. 3; Se-281;
Matches 765; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 606 MFHAPLTINPLCKVIRFNIDYTHFIEEMMPENFCVKLELFSLEFIDILEYDMNLK 665
Db 1 MFHAPLTINPLCKVIRFNIDYTHFIEEMMPENFCVKLELFSLEFIDILEYDMNLK 60
QY 666 GPLFEDSPCCPRFHPMPRVRLPDGKEVLSMHQILLYLRCSALVPEEELANMLQW 725
Db 61 GPLFEDSPCCPRFHPMPRVRLPDGKEVLSMHQILLYLRCSALVPEEELANMLQW 120
QY 726 EELEWOKYAECKGMIVTNPGRKPSVSRIDQIDREQFNPDVITFPIIVHGRPAQLSTA 785
Db 121 EELEWOKYAECKGMIVTNPGRKPSVSRIDQIDREQFNPDVITFPIIVHGRPAQLSTA 180
QY 786 GDPYOGLMKSYVYKLNHLANSPPKVQTDOKIQAQREALOKIROKNTMRRETVELSSQ 845

```


Db 181 GDFQYQKLKSKYVLRHLLANSBPYKQTDKQKLAOREALQKIRKNTMRREVTVELSSQ 240
QY 846 GFWKTIQRSDVCOHAMMLPYLTHHRYHOCMLHDLKIGTYFQDRCLQTLATHTSHHNL 905
Db 241 GFWKTIQRSDVCOHAMMLPYLTHHRYHOCMLHDLKIGTYFQDRCLQTLATHTSHHNL 300
QY 906 FGMNPDHARNSLSCGIRQPKYGRKRVHMHMKKGINLIMSRGODDPTPSKRNIN 965
Db 301 FGMNPDHARNSLSCGIRQPKYGRKRVHMHMKKGINLIMSRGODDPTPSKRNIN 360
QY 966 ERLEFEGDAVEFLTVSHLYLTPSPSEEGCLATTAIVQONHMLAKKLEIDPMMLYA 1025
Db 361 ERLEFEGDAVEFLTVSHLYLTPSPSEEGCLATTAIVQONHMLAKKLEIDPMMLYA 420
QY 1026 HGFDLCRESDLRHAMNCEBALGAVYLKSGSEEAQKLGRLFPNDPDLREVMANPLHP 1085
Db 421 HGFDLCRESDLRHAMNCEBALGAVYLKSGSEEAQKLGRLFPNDPDLREVMANPLHP 480
QY 1086 LQLOEPNTDROLIETSPVLQKLTPEEBAIGVITFHVRLRLAFTLTGFMHLLTGHNR 1145
Db 481 LQLOEPNTDROLIETSPVLQKLTPEEBAIGVITFHVRLRLAFTLTGFMHLLTGHNR 540
QY 1146 MEFLGDSIMQVATEYLTFHPDHEGHGHTLRLSSLVNNRQAKAAEELGMOEYATINDK 1205
Db 541 MEFLGDSIMQVATEYLTFHPDHEGHGHTLRLSSLVNNRQAKAAEELGMOEYATINDK 600
QY 1206 TKRPVGLRTKTADLLESFIALYTDKDEYVTFEMNVCFFPLKFEILNODMNDPKSQL 1265
Db 601 TKRPVGLRTKTADLLESFIALYTDKDEYVTFEMNVCFFPLKFEILNODMNDPKSQL 660
QY 1266 QQCCLTLRTGKRPDIPLKTLQTVGSPSHARTYVAVYFGERIGCGKSPSIQQAEMGAA 1325
Db 661 QQCCLTLRTGKRPDIPLKTLQTVGSPSHARTYVAVYFGERIGCGKSPSIQQAEMGAA 720
QY 1326 MDLLEYNFPQAHQKRFGRKTRQDLAKEMREREQEPDETEIKK 1374
Db 721 MDLLEYNFPQAHQKRFGRKTRQDLAKEMREREQEPDETEIKK 769
RESULT 3
Q9XYNS PRELIMINARY: PRT: 1327 AA.
ID Q9XYNS
AC Q9XYNS:
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE DROSHA PROTEIN.
GN DROSHA OR CG8730.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Goebye J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Ball J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,
RA Burits K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Donip L.E., Donnes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
RA Jajall M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Pan S., Pollard J., Puri V., Reese M.G.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shee B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Splyer E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Sprakes R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Filippov V.A., Filippova M.A., Gill S.S.;
RT "The complete sequence of Drosophila droscha gene that contain two
RT Rnaase III signatures.";
RL Submitted (DRC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE003839; AAF59169.1; -
DR EMBL: AF116572; AAD51170.1; -
DR FlyBase: FBgn0026722; droscha.
DR InterPro: IPR001159; DS_RBD.
DR InterPro: IPR000999; Rnaase_3.
DR Pfam: PF00035; dnm_1.
DR Pfam: PF00636; Ribonuclease_3; 2.
DR SMART: SM00358; DSRM; 1.
DR SMART: SM00358; RIBOC; 2.
DR PROSITE: PS0137; DS_RBD; 1.
DR PROSITE: PS00517; Rnaase_3; 1; 1.
DR PROSITE: PS01442; Rnaase_3; 2; 2.
SQ SEQUENCE 1327 AA; 152816 MW; 84AD3AABD676C9E8 CRC64;
Query Match 39.5%; Score 2963; DB 5; Length 1327;
Best Local Similarity 45.6%; Pred. No. 5.4e-202;
Matches 621; Conservative 213; Mismatches 327; Indels 200; Gaps 32;
QY 90 GGPLRPPCFIRPPFPNHQKRPVPPPCFPMPMPMPPCPN-NPPVPGAP-----P 137
Db 3 GGPLRPPPVQPA-----PPPPPPPEEDLSPGVGVSHNYSNESHQ 46
QY 138 GCGTFFPMKPPRSMRPPPPVMPQVNYGYRPGY-----SHHNP 179
Db 47 SKSLDYVYPRTPADYASS---VPSYDYPQDPAYGEGYAVNEQAKYGGQESHYQY 103
QY 180 PSFNSFQNNPSSFLPSANSSSPHRLPLPYPL-----PKASERSRPERLKHYYDHRND 235
Db 104 PA-----SGSFLVES-----YKYPDRYPVASSNVRPSEHQ-----RYTS 139
QY 236 KSHGCGEHRSLDRKRGSRPDRRQDSRYSDVDKGTBRHSHYSERSRREERHNR 295
Db 140 NSSSGYHHYP-----GYSGSR-----RYDGRHDO-----EHRQIDOSRANHPRGHY 183
QY 296 DNRASPLERSYKKKRYKSGRSYGLSVVPEDACGPPELPGELIKNTDSMAPLEIVHRS 355
Db 184 AHRQKSGSHGY--YGSAAKN-----QVSDYSFRGHHEERN 219
QY 356 PSREKKRA--RWEKKD--WSDNOSGDKNYSITKEKEPETMPPDKNEEBELK 409
Db 220 FLEKTRAKPVEYERDRDLRWQCSNFC-----EKPDYKKMNALEADAPV 267
QY 410 PYWICHTSENYSSDDPDQYGDSTVGTSLRLDYKFEELCSROBKAKAAPRPEPP 469
Db 268 ESWVSSPAELIYERTK-----SENEVGRARLQKTLCTFDELLQRAKRVREKLPVYVP 323


```

Db 242 EDLESSESECESDSDSTCSSSDSEVFYIAIEIKKKAKHAPRLDAHELYNDPGQMDGP 301
QY 535 LCKCSAKKARTGIRHSITYGGEAICPCRPMTNNAGRLPHYRTTVAPPNFLDRVLEY 594
Db 302 LCKCSAKKARTGIRHSITYGGEAICPCRPMTNNAGRLPHYRTTVAPPNFLDRVLEY 361
QY 595 DDHEYTFEGFSMFANAPLT 613
Db 362 DDHEYTFEGFSMFANAPLT 380

RESULT 7
001327 PRELIMINARY; PRT; 1086 AA.
AC 001327;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
GN F26B4.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN (1)
RP SEQUENCE FROM N.A.
RA Lightning J.;
RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RT none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; 281070; CAB03006.3; .
DR InterPro; IPR001159; DS_RBD.
DR Pfam; PF00035; dsrm. 1.
DR SMART; SM00358; Rbionuclease_3; 2.
DR SMART; SM00353; RIBOC; 2.
DR PROSITE; PS50137; DS_RBD; 1.
DR PROSITE; PS00517; RNASE_3_1; 2.
DR PROSITE; PS50142; RNASE_3_2; 2.
SQ SEQUENCE 1086 AA; 125333 MW; 4A478120F88F8PB8 CRC64;

Query Match 19.3%; Score 1448; DB 5; Length 1086;
Best Local Similarity 31.2%; Pred. No. 2.7e-94;
Matches 364; Conservative 206; Mismatches 431; Indels 166; Gaps 31;

QY 277 RHRSESRERERHRHRNDNRSPSLSESYKKEYKRSGRSYGSLVPEPAGCPPELP-G 335
Db 14 KHKRRARRKKYQAEYQBRKHEEMQOLGRFQNPSTSSAPPTVYKLTLPRESTSALPFG 73
QY 336 EIKKNTDS-----WAPPLEIVNHRSPSRKKRAWEEDKDRMSQSGKDKNTYSIK 388
Db 74 DSPRLTEKYETNYMIDPP--VYSHSAELIK-----SNRYVIK 110
QY 389 EKEPEETMPDKNEEEELK-----PWTIRCHS----- 418
Db 111 ABEAKRYMMIKKSTTSTLQDEQKILLETYKTKRRLOADVPIIIPCHSMKGRKTPKQK 170
QY 419 ---ENYSSDPMDQVGDSTVVTSLRLDLVDFEELGSRQEKAKARP--WEPPKTKL 473
Db 171 GGDEFTASDVSDSDSND-----QDEASTSEPTTRQAPLEADKT 208
QY 474 DEDLESSSECESEDSTCS-----SSDSEVFYIAIEIKKKAKHAPRLH 519
Db 209 GR-----VKDEKQTCNRRNOQRAKRLRNFEKEKQITLLKKGIDRKKTHPNQIH 258
QY 520 DELMYNDPGQMDGPLCKCSAKKARTGIRHSITYGGEAICPCRPMTNNAGRLPHY--RIT 577

```

```

Db 259 PDISEFNKGLGNEBPECRCPEPIKTGLKHYGAYAGEDKAIQK--KSNGENLHYTLRVAT 316
QY 578 VSPPTNLTDRPVYIEEDHEITYFEGFSMFANAPL-----TNPLCVYIRNDITYTHFIE 633
Db 317 PLPSENOL--YRTHMALNGEEFEFEGRSLTTHAPLPDCMRADIC--YSMDYEQLE 371
QY 634 EMMP--ENFCVAGLEFLSLFRLIELYDWNLKGPLFEDSPCCPFFHMFPRFVRLEPVG 692
Db 372 EFMFDECFDEPDCDMLFEYTFHELFEMLDRLKPHIPSVYECSPMHIIMPRVQ--TKDD 430
QY 693 GKEVLSMHQILLYLL--RCSKALVPEEELIANMLQWEELEMQVAEECKGMLVTNPGKPS 751
Db 431 LVQLMSKTVLAFTSKSGSEIMSPEDVNRICDAQIDQFTFRNSKHOSIVLTNTEKPESA 490
QY 752 VRIDQLRDEQPNPDVITFPIIVHNGITRPADLSVAGSDPOYQKMKSVYKLRLHLNAPYK 811
Db 491 IRADWFERDEKKEV---YVHNALRAQTYTAISLPRIAFLEKTYLKMIO--EKSSGVY 545
QY 812 QTDKOKLAOREEALOKIRQKNTMRREVTVELSSQGEFWKTGIRSDVCQHAMMLPVLNHR 871
Db 546 NKDEFK---TKNELEHLKRENRBSARNIKLEPYAGFTIENGKLDVAHVMTTLACHHR 602
QY 872 YHOCMLMLDKLIGTFODRCLQLAMTHPSHILNFGMNPDHARNSLSCGTRQPKYG--D 929
Db 603 YNFSLDVFEERYIEYKFNDRRVIELALMHSFSGHYGTPIDVHKNMTNCGYRR--KYGARD 661
QY 930 RKVYHMMRRKKGINTLINISRLGODDPRPSRINNHRELEFGDAVEFLTVHLYLPP 989
Db 662 KR---EKKRVAGIWSLRIKNGKTSGGP---TLHNERLEYLDGDAVELVSHHLYFMLT 714
QY 990 SLEBGLATYRTAIVQNONHLAMLAKKLELDPFMIYAHGPDLCRESDLRRAMANCPEALIG 1049
Db 715 HHFEGGLATYRTALVQNRNLATLAKNCRIDEMLYSHGADLIIVAEFKHALANAPEVAVA 774
QY 1050 AVYIEGSLLEAKQLFGKLF--NDPDLREWYLNTPRLPLQLOEPNTRQOLIESPVYLQKIT 1108
Db 775 AIYLDGLIAPCDVIFSKAMYGHPVLKEKWDHINHEHLKREDQGDRODSFTRPSTIEPH 834
QY 1109 EFEELGIVTFHVLARATFTYVGNHNLTLGHNOBMELFGLDAVEFLTVHLYLPP 1168
Db 835 ALERRLGIQNNIRLKLAKATTRRNIPNNDLTGKHNQLELGSVQLLVSDLEIRREY 894
QY 1169 HHGHLTLRSLVNNNTQAKVAEELGMOEYATINDTKRPV---GIKRTYLADELSEPI 1225
Db 895 HHGHSMLTSLVSNQTOAVVDDLGFTFVY-----KAPYTPRLKDKRADIVLEAFI 949
QY 1226 AALYTDKDELYVATFMNVCFPRFKERFILNQDNNDRKSQLOCCCLFLR--TEGKEPDIPL 1284
Db 950 GALYVDGIGHCRAFIYIVCPRIKHFIESEKYNDAKSHLOQCLAMRDPSSSEPDMPY 1009
QY 1285 KTLQTVGSHARTYTAIVVYKGERIGCGKPSIQOAEQMAAMALKYNPPOAHQKRF 1344
Db 1010 RVLGIEPTNNRIFKTAIVYIKGKRILASAESVNHAKLELYAELALA--NLEKSPSK--- 1064
QY 1345 GRKYRQELKEMREHOREPDETD 1371
Db 1065 -----MKAKNNSWFQNNRRRLQDTS 1086

RESULT 8
090908 PRELIMINARY; PRT; 860 AA.
AC 090908;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
GN RIBONUCLEASE (FRAGMENT).
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN (1)

```

RP SEQUENCE FROM N.A.
 RC STRAIN-CH1489 HIM-8(E1489);
 RX MEDLINE-20179701; PubMed-10713462;
 RA Filippov V., Solovyev V., Filippova M., Gill S.S.;
 RT "A novel type of RNase III family proteins in eukaryotes."
 RL Gene 245:213-221(2000).
 DR EMBL: AF160248; AAD45518.1; -
 DR InterPro: IPR001159; DS_RBD.
 DR InterPro: IPR000999; RNase_3.
 DR Pfam: PF00035; dsrm; 1.
 DR Pfam: PF00636; Ribonuclease_3; 2.
 DR SMART: SM00358; DSRM; 1.
 DR SMART: SM00355; RIBOC; 2.
 DR PROSITE: PS00137; DS_RBD; 1.
 DR PROSITE: PS00517; RNase_3.1; 2.
 DR PROSITE: PS0142; RNase_3.2; 2.
 FT NON_TER
 SQ SEQUENCE 860 AA; 99414 MW; B339A3E7C9622A08 CRC64;

Query Match 18.9%; Score 1417; DB 5; Length 860;
 Best Local Similarity 36.2%; Pred. No. 3, 2e-92;
 Matches 318; Conservative 173; Mismatches 334; Indels 54; Gaps 21;

QY 508 IKRKAHPRLDELWYNDPGQMNQDPLCKCSAKARRTGRHSIYGEERAIKPCRPMTNN 567
 DB 21 IDRRKTHPNCIHPDISFENEGKLNCEGECRCPEPIKCGLKHGYAGEDKADICK--KSN 78
 QY 568 AGRLEFH--RTVSPPTNLTDRPYVIEDHDYEIEGEGSMFNAAPL---TNPICLKYI 621
 DB 79 GENHLYTLVNTPLSPENQ--YRTNHAINGEEFEFGSLTHNAPDCMTAPKIC-- 134
 QY 622 RFNIDYTHRIEEMMP-ENFCVKGLELSFLRDLILELDMNLKGPLFEDSPCCPRFH 680
 DB 135 -YSDDYEQVLEERMPDECPDPCDMLFEYIFHEIFEMDFELRPHHPSDVESCPMH 193
 QY 681 EMRFVAFRLPDGKEVLSMHOILLYL-RCKSLVPEEETANMLQWEELEMQRYAECKG 739
 DB 194 IMPRFVQ-TRDDLVQMLSSKTVLAYFTSKSSSEIMSPEDVNRCLDQIDQFTNTSKHK 252
 QY 740 MIVNPTCKSSVRIOLDREOFNPVITPPIVHGIRPAQSVAGDPQYQKLMKSYK 799
 DB 253 SYLVNTFKFSATRADFERDEEKEY---YVHNAIRQYTTAISLPRIAFLEKTLNK 308
 QY 800 LRHLANSPVKQTDOKLAQREBALOKIKQKTMREYVELSOGFMWTKGRSDYCOH 859
 DB 309 MIQ-EKQSSGVYKNDKFEK--TKNELEHLKREKRSANLRLRPVAGFITGLKPRVAAH 364
 QY 860 AMMLPVTTHIRYHOCMLDKLIGYFQDRCLLQTLAMTHPSHNLNFGMNPDRARNSLN 919
 DB 365 VVMTILACHHIRYNFSLDVEEYEVYKFNDRVYELALMHSFKSHYGPIDHVKMTIN 424
 QY 920 CGTROPKY--DRKVHNMHRKKGINFLIMSLGDDTPSRINNELEFLGDAVVE 977
 DB 425 CGTNR-KYGAEDKR--EKKRVAGIMFLNMGTSGER---LHNELEVLGDAVVE 476
 QY 978 FLTSVHLUYFSPLEEGGLATYRTAIVQNOHNLAMAKLELDPMLYAHSPDLCRESDLR 1037
 DB 477 LYSVHNLUYMLTHFEGGLATYRTALQVONNLTATLAKNCIDELQYSHADLLINAEFK 536
 QY 1038 HAMANCEBALIGAVLEGSLEAKQFGRLLF-NDPDLREVLWLYRPLHQLQDEPNTDQ 1096
 DB 537 HALANAEAVMAALYLDGLGAPCDVIFSKAMYGHOPLAKEMWHINHELEKREDPOGDRD 596
 QY 1097 LIERSPLQKLEFEELAGIYFTHVRLARAFILRTYGFNHLTIGHNQRKEFLGDSIMOL 1156
 DB 597 LSFITPLTFHFALDEERLQGFQFNIRLLAKAFTRKRNIPNDLTKGHQRLWLEMGDSVLQ 656
 QY 1157 VATEYLFTHRPDHEGLTLTSSLVNRRTOAKVAEELAQGEAYITNDKRRPV---GLR 1213
 DB 657 IVSDFLVRFRYHHEGMSLRTSLVSNOTQAVVCDLGLTEFY-----KAYKPELAK 711
 QY 1214 TKTLADLESFIALYDQKDELYVHTFMVCFPRLEKEFLINDQMDPKSQLOQCCLTLR 1273

DB 712 LKQADLVEAFIGALYVDRGLEHCRAFLIRVCPRLKHFISEKWNDAKSHLQWCLAMR 771
 QY 1274 -TEGKPPDPLKTTQTVGSPSHARTTYAVVEFKRGCGKSPISQAEAGMADALEKY 1332
 DB 772 DSSSEPPMPPEYRVLCIEPTNNRIFKTAIVYKKGRLASAESNVHKAELVLAELA-- 829
 QY 1333 NFPQAHQKRFIGRRYROELKEMWREHREHREPPED 1371
 DB 830 NLEMSFSK-----MKAKNSWFOUMRRRLQDTS 860

RESULT 9

QYCTG2 PRELIMINARY; PRT; 134 AA.

AC 09CTG2;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE 1110013A17RIK PROTEIN (FRAGMENT).
 GN 1110013A17RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6J; TISSUE-EMBRYO;
 RX MEDLINE-21085660; PubMed-11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
 RA Fiedlschmann W., Gaasterland T., Gissi C., King B., Kochava H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guslinich S., Hill D., Hofmann M., Hume D.A., Kamlay M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Morone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyszaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK003651; BAB22917.1; -
 DR MGD: MGI:1915895; 1110013A17RIK.
 DR InterPro: IPR001159; DS_RBD.
 DR Pfam: PF00035; dsrm; 1.
 DR SMART: SM00358; DSRM; 1.
 DR PROSITE: PS00137; DS_RBD; 1.
 FT NON_TER
 SQ SEQUENCE 134 AA; 15883 MW; 72AA673BFDCA8A1C6 CRC64;

Query Match 9.5%; Score 711; DB 11; Length 134;
 Best Local Similarity 97.8%; Pred. No. 4, 1e-43;
 Matches 131; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1241 MNVCFPRLEKEFLINDQMDPKSQLOQCCLTLRTGKREPDILPLYTLQTVGSPSHARTTYV 1300
 DB 1 MNVCFPRLEKEFLINDQMDPKSQLOQCCLTLRTGKREPDILPLYTLQTVGSPSHARTTYV 60
 QY 1301 AVEFKERIGCGKSPISQAEAGMADALEKYNFPQAHQKRFIGRRYROELKEMWRE 1360
 DB 61 AVEFKERIGCGKSPISQAEAGMADALEKYNFPQAHQKRFIRKRYROELKEMWRE 120
 QY 1361 HOEREPDEDEDIKK 1374
 DB 121 HOEREPDEDEDIKK 134

```

RESULT 10
OSVAV4 PRELIMINARY: PRT: 1151 AA.
ID OSVAV4
AC OSVAV4:
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE CG5514 PROTEIN.
GN CG5514.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Brokstein P., Brotlier P.,
RA Borova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jaitli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laspo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maltel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merulov G., Mishina N.Y., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Rehert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinslock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Gibbs R.A., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003763; AAF56763.1; -.
DR Flybase: FBgn0039560; CG5514.
DR InterPro: IPR002965; P_richextensn.
DR PRINTS: PR01217; PRICHEXTENSN.
SQ SEQUENCE 1151 AA; 125450 MW; AF8330378658C1D1 CRC64;

Query Match 4.5%; Score 340; DB 5; Length 1151;
Best Local Similarity 22.5%; Pred. No. 2,1e-15;
Matches 173; Conservative 83; Mismatches 280; Indels 234; Gaps 35;

```

```

Db 416 PPAPPTIKPPPPAPPTVEPPP-----PPAPPTVEPPPPPPAPPTKVEPPPPAPAE 469
Qy 137 --PGQTFPPMPMPSPMHP-----PPVMPQOVNTQYPP-----GY----- 172
Db 470 VEPPPAPPTLEPPPPPPAPPKVELPPPPAPPAEAIAITRRKAGSSFTTELAATPKES 539
Qy 173 ---SHNFE-----PPSFNSFQNP-----SFLPSANSSSPH 203
Db 530 ETRKPNHNVTDYKEDQPAETEDKQKPLEGLDAPKDAVSTASSTESTSTSSOS-- 587
Qy 204 FRHLPPPLKAPBESRPERLKHVDHRDRHDSHGEGHRLDR-REGRSPDDRROD 262
Db 588 -----KSHSKSDKKEKDRRHRHSDKH-----RRSTDHSDRSDKSHSKS 633
Qy 263 SRRYSDYDRG-PPSRHSYRSRERERHRHRDRNRSPSLR-----SYRK 309
Db 634 SSSSSSSSSSSSSSSSKKSSSSSKNDKSSSSSSSSSSSSSSSSSSSSSSSSSSSR 693
Qy 310 -----EYKRSRYSGLSVPEPAGCTPELPGETIKNTDSWAPPLEIYNHRSRERKAR 364
Db 694 HKSSSSSSSRSEBDKG-----KEKDKEREKDSQRS-----HSSSSSSSSSR 736
Qy 365 WEEKDKWD-----NOSGKDN-----YTSIE-----KEPETMDKNE 401
Db 737 KDHRGRDRNRKNSNTSGSENKAIHDDHSESEKRYQRGRSDSNDGKPPSGGPAKNS 796
Qy 402 EEEDE--LKPWIRCTHSENYYSDPMQVDS---TVGTSRLDYDKFEELGSR 455
Db 797 QPESAAITMSDAVEANANTNGNSNGTACADNVGIVYDIIQSTSSVELTAAS 856
Qy 456 QEKARAPPEWEPKTKLDELESSSECESEDSTCSSSD-----SEFPDVAIEIK 509
Db 857 QSHRAASSKHEPEDIDGKEADNQPEKTELEARDQ-CASQNEVPIVPEQTLADSVDP 915
Qy 510 RKAHPRLDLMLYNDPGOMNGPLCKCAKARIGIRISITPGREATKPCAPMNNMG 569
Db 916 AKES-MTLEDE-----KVTVEBN--KSEESKP----- 943
Qy 570 RLHYRTVSPPTNLTDRPTVLEYDDHEYIFGFSMAFAPLTNPLC--KYIRPNIDY 627
Db 944 -----ENPPECVDEPAGQVDPE-----PQADKPKATPVPISEQSDERSADF 988
Qy 628 TIFTEEMPEPNCVKGLEFLFLFRDILLYDW-----NKGLEFED 671
Db 989 VTHF-EENTDE-----FTRRLQNLINOLEDKKNLNLRLSD 1023

RESULT 11
ID Q16630 PRELIMINARY: PRT: 551 AA.
AC Q16630;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE HBBRII-4 MRNA.
GN HBBRII-4 OR HBBRII-7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fleischhauer K.L.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: X67337; CAA47752.1; -.
DR EMBL: X67336; CAA47751.1; -.
DR InterPro: IPR000504; RRM.
DR Pfam: PF00076; rrm; 1.
DR SMART: SM00360; RRM; 1.
DR PROSITE: PS5102; RRM; 1.
SQ SEQUENCE 551 AA; 59208 MW; 407CB941BA56A966 CRC64;

Query Match 4.2%; Score 316; DB 4; Length 551;

```



```

Db      889  TPKAKSHSKAKEHOEAKPAKDEKVKKDKCKDIKSEKPAK--DEKAKKPKKNLDSKGEK  947
QY      545  -----TGIRHSITYPEEAIKPC--RPMNNMGRLEH-----YKITYSP  580
Db      948  RKRKTEEKVDKDFESSKISKKECTELVKPSPKRMKGDEKLERPEKDKIASSTTP  1007
QY      581  PTFNLTDRPTVLEYDDEHETFEFGFSMFAHAPLNIPLCKVIFNIDYTHFTEEMMPENF  640
Db      1008 AKKIKINRETGGKIGNAE-----NASTTKPSEK-----LESTISKIKÖEKVKK  1052
QY      641  C---VKGLEFLSLFLFRDILLEYDMNLKGLFEDSPCCRFHFMPRFVRLPDGKEVL  697
Db      1053 AKRKVAGSESSSTLY-----DYSTSGSGSPVR  1082
QY      698  SMHQLLYLIRCSKAL-----VPEERIANMLQ-----WEELEOKAYAECK-----  738
Db      1083 KSEKTDTRKRYAIKIMEEYNNNDTAPAEDEVITIMIQPSKMDKDDFESSEEDVKTTOPIQ  1142
QY      739  -----GMIVNFGTKPSSYRIDQLDREQFNPDVITFPPIVHGIRPALSYAGDPOYOKL  793
Db      1143 SVGKPSIITIKNTYTKPSAT-AKYTEKESQPE-----KLOKL  1178
QY      794  WK--SYVKLRHLI-----ANSPKYQTDKQKLAOREALQIRQKNTMRREYVE-LSS  844
Db      1179 PKASHELMQHELRSSKGSASSEKGRAKDREHSGSEKNDPKRKSQAQDPKESTVDRLSE  1238
QY      845  OGFWKT 850
Db      1239 OGFWKT 1244

```

RESULT 15

```

Q9SBM1
ID      Q9SBM1      PRELIMINARY:      PRT:      409 AA.
AC      Q9SBM1;
DT      01-MAY-2000 (Tremblrel. 13, Created)
DT      01-MAY-2000 (Tremblrel. 13, last sequence update)
DT      01-JUN-2001 (Tremblrel. 17, last annotation update)
DE      HYDROXYPROLINE-RICH GLYCOPROTEIN DZ-HRGP PRECURSOR.
GN      HRGP GENE.
OS      Volvox carterl f. nagariensis.
OC      Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OX      Volvocaceae; Volvox.
RX      NCBI_TaxID=3068;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=HK10;
RX      MEDLINE=20044763; PubMed=10574980;
RA      Ender F., Hallmann A., Amon P., Sumper M.;
RT      *Response to the sexual pheromone and wounding in the green alga
RT      Volvox: Induction of an extracellular glycoprotein consisting almost
RT      exclusively of hydroxyproline.
RL      J. Biol. Chem. 274:35023-35028(1999).
DR      EMBL; AJ242540; CAB62280.1; -.
DR      InterPro: IPR003882; Pstl1l_extensin.
DR      InterPro: IPR002965; P-rich_extensn.
DR      PRINTS; PR01217; PRICHEXTENS.
DR      PRINTS; PR01218; PSTLEXTENSIN.
KW      Signal.
FT      SIGNAL.
FT      CHAIN.
SQ      SEQUENCE      409 AA: 41547 MW: 40074906AF02BD74 CRC64,

```

Query Match

Best Local Similarity 3.8%; Score 285; DB 10; Length 409;

Matches 76; Conservative 11; Mismatches 79; Indels 32; Gaps 6;

```

QY      31  APFRRQNLRLHPQRPVQYQYEPSSA--PSTTFSNAPNPLPRPRDF-----VPFPP  84
Db      19  APARKRP-----PRRSVALVETPAAPRGSPRGTPPGVPRPTPSGPHRPPPP  72
QY      85  MPSSAGPLRPCPIRPPFNHQRHPRFVVRPCFPRMPCRNPRVPGARRGGTFPF  144

```

```

Db      73  PPPPQPLPPSPSPPPPP-----PPVPPPPPPPPPPPPSPPPPPPP-----  119
QY      145  MMPSPSMHPPPPPVVMQOVNTQYTPPGISHHNFPPSPFSNQNNPSSFLSANNSSPHF  204
Db      120  --PPSPSPPPPPPPPPPPPPPPPPPP-----PPSPSPSPSPSPSPSPSPSPSP  172
QY      205  RHLPPYLPKAPSERRSP 222
Db      173  PPSPPSPSPSPPPPPPP 190

```

Search completed: October 9, 2002, 18:25:49
Job time : 100 secs

